



TITLE:

Optimization algorithms on the Grassmann manifold with application to matrix eigenvalue problems

AUTHOR(S):

Sato, Hiroyuki; Iwai, Toshihiro

CITATION:

Sato, Hiroyuki ...[et al]. Optimization algorithms on the Grassmann manifold with application to matrix eigenvalue problems. Japan Journal of Industrial and Applied Mathematics 2014, 31(2): 355-400

ISSUE DATE:

2014-06

URL:

<http://hdl.handle.net/2433/199668>

RIGHT:

The final publication is available at Springer via <http://dx.doi.org/10.1007/s13160-014-0141-9>; This is not the published version. Please cite only the published version.; この論文は出版社版ではありません。引用の際には出版社版をご確認ください。

Noname manuscript No.
(will be inserted by the editor)

Optimization algorithms on the Grassmann manifold with application to matrix eigenvalue problems

Hiroyuki Sato · Toshihiro Iwai

Received: date / Accepted: date

Abstract This article deals with the Grassmann manifold as a submanifold of the matrix Euclidean space, that is, as the set of all orthogonal projection matrices of constant rank, and sets up several optimization algorithms in terms of such matrices.

Interest will center on the steepest descent and Newton's methods together with applications to matrix eigenvalue problems. It is shown that Newton's equation in the proposed Newton's method applied to the Rayleigh quotient minimization problem takes the form of a Lyapunov equation, for which an existing efficient algorithm can be applied, and thereby the present Newton's method works efficiently. It is also shown that in case of degenerate eigenvalues the optimal solutions form a submanifold diffeomorphic to a Grassmann manifold of lower dimension. Furthermore, to generate globally converging sequences, this article provides a hybrid method composed of the steepest descent and Newton's methods on the Grassmann manifold together with convergence analysis.

Keywords Grassmann manifold · Riemannian optimization · Steepest descent method · Newton's method · Rayleigh quotient · Lyapunov equation

Mathematics Subject Classification (2000) 65F15 · 65K05 · 49M15

1 Introduction

A problem of minimizing the value of a given function subject to some constraint conditions is called an optimization problem. A number of methods have been developed to solve optimization problems on the Euclidean space. For a constrained problem on the Euclidean space, however, if the set of points which satisfy constraints forms a smooth manifold, the problem can be considered as an unconstrained problem on this manifold. In their paper entitled "The geometry of algorithms with orthogonality constraints," Edelman, Arias, and Smith developed

H. Sato (Corresponding author) · T. Iwai
Department of Applied Mathematics and Physics, Graduate School of Informatics,
Kyoto University, Kyoto 606-8501, Japan
E-mail: hsato@amp.i.kyoto-u.ac.jp

several algorithms for unconstrained problems on the Stiefel manifold and on the Grassmann manifold [6]. In their book entitled “Optimization Algorithms on Matrix Manifolds,” Absil, Mahony, and Sepulchre developed algorithms for optimization problems on a general Riemannian manifold and discussed the convergence properties of the algorithms [2].

In an optimization problem on the Euclidean space, the simplest way to compute the next iterate is to perform a line search along a search direction. In a problem on a manifold M , however, once a search direction is given as a tangent vector to the manifold M , one needs an appropriate map from the tangent bundle TM to M in order to determine the next iterate. Such a map is called a retraction, which is one of the most important concepts in optimization algorithms on manifolds, and effective use is made of geometric tools for defining such a map. The retraction maps a tangent vector onto a geometrically appropriate point on M , which makes the optimization algorithm accurate. For constraint problems on the Euclidean space, the gradient projection method [15] works well if the constraint is piecewise linear, but it could not be effective for generic constraints, since the counterpart of the retraction is not exactly defined.

In [6] and [2], the Grassmann manifold is treated as a quotient manifold of the matrix Euclidean space. However, the present article deals with the Grassmann manifold as the set of all orthogonal projection matrices of constant rank and sets up several algorithms in terms of such matrices. The same view has been taken in [12], but the approach to optimization problems taken in this article is different from those in [12] on the point that the optimization algorithms in this article make use of the Stiefel manifold which projects to the Grassmann manifold. More specifically, for a point X of the Grassmann manifold, there exists a point Y in the Stiefel manifold such that $X = YY^T$, which is utilized in our optimization algorithms. This makes differences between the existing and the proposed algorithms in their details. Apart from optimization problems, the Grassmann manifold is used in physical problems such as coherent states and quantum computers, in which the Grassmann manifold is treated as the set of Hermitian projection matrices of constant rank (see [8] and [16], for example). Further, the Stiefel and the Grassmann manifolds provide interesting aspects on optimization and control problems [11].

The organization of this article is as follows: The geometry of the Grassmann manifold $\text{Grass}(p, n)$ as the set of all $n \times n$ orthogonal projection matrices of rank p is discussed in Section 2 together with the Stiefel manifold $\text{St}(p, n)$, where $\text{St}(p, n)$ is the set of all $n \times p$ orthonormal matrices and projected to express elements of the Grassmann manifold $\text{Grass}(p, n)$. Materials needed to describe algorithms, such as geodesics, the gradient and the Hessian of a function, *et al.*, on the Grassmann manifold are set up in terms of orthogonal projection matrices of rank p . In Section 3, after reviewing the steepest descent method and Newton’s method on a Riemannian manifold, the corresponding algorithms on the Grassmann manifold are developed in terms of orthogonal projection matrices of constant rank together with applications. An algorithm is also provided to compute $Y \in \text{St}(p, n)$ such that $X = YY^T$ for a given $X \in \text{Grass}(p, n)$. The Rayleigh quotient cost function, which is typically treated on the Stiefel manifold, reduces to a function on the Grassmann manifold because of symmetry. The resultant function is called the reduced Rayleigh quotient cost function, which is occasionally referred to as the Rayleigh quotient for short. Newton’s equation for the Rayleigh quotient on

the Grassmann manifold is shown to be expressed as a Lyapunov equation, and can be solved by applying an existing algorithm. Further, an intensive analysis is made of the reduced Rayleigh quotient cost function with insight into the Hessian at critical points, and thereby a global optimal solution is characterized. In particular, it is shown that if the symmetric matrix characteristic of the Rayleigh quotient cost function has a degenerate eigenvalue of particular type, the set of global optimal solutions for the reduced Rayleigh quotient cost function forms a submanifold diffeomorphic to a Grassmann manifold of lower dimension and the Hessian is degenerate on this submanifold. In Section 4, issues encountered in applying solely Newton's method are addressed first, and then a hybrid algorithm composed of the steepest descent method and Newton's method is provided, which is shown to exhibit good convergence property together with numerical experiments performed for the reduced Rayleigh quotient cost function. In addition, it is shown that Newton's method serves to improve an approximate solution to a matrix eigenvalue problem. For example, eigenvalues obtained by MATLAB can be improved. In fact, the value of the Rayleigh quotient for the solution obtained by Newton's method is less than the value for the solution by MATLAB. Furthermore, a question as to whether generated sequences converge to a global optimal solution or not is examined, which is encountered in the application of the steepest descent method. Sufficient conditions are given for a sequence generated by the steepest descent algorithm to converge to a saddle point. However, the set of those initial points with which sequences generated by the algorithm start to converge to saddle points is of measure zero. Moreover, even if the initial point is put on this measure zero set within the computer accuracy, the limitation of computational accuracy brings the generated sequence out of this set, and thereby the sequence converges to a global optimal solution, practically. In Section 5, this article concludes with some remarks about the difference between some preceding studies and the present results.

2 The geometry of the Grassmann manifold

Several results stated in this section may follow from [12] or [6], but our approach is different from theirs. We make remarks in place for comparison with them.

2.1 Definitions

Let n and p be positive integers with $n \geq p$. Let $\text{Grass}(p, n)$ and $\text{St}(p, n)$ denote the set of all p -dimensional linear subspaces of \mathbb{R}^n , and the set of all $n \times p$ orthonormal matrices, respectively, where $\text{Grass}(p, n)$ and $\text{St}(p, n)$ are called the Grassmann and the Stiefel manifolds, respectively [2, 3, 6, 7, 18]. It is well known that the Grassmann manifold is a factor space of $\text{St}(p, n)$;

$$\text{Grass}(p, n) \simeq \text{St}(p, n)/O(p), \quad (2.1)$$

and of $\mathbb{R}_*^{n \times p}$;

$$\text{Grass}(p, n) \simeq \mathbb{R}_*^{n \times p}/GL(p), \quad (2.2)$$

where $\mathbb{R}_*^{n \times p}$ is the set of all $n \times p$ matrices of full-rank. As is easily seen from (2.2), the dimension of $\text{Grass}(p, n)$ is

$$\dim(\text{Grass}(p, n)) = p(n - p). \quad (2.3)$$

In [6] and [2], the Grassmann manifold $\text{Grass}(p, n)$ is treated on the definitions (2.1) and (2.2), respectively, and several algorithms on $\text{Grass}(p, n)$ have been developed with these definitions.

The Grassmann manifold $\text{Grass}(p, n)$ can be, however, viewed as the set of all orthogonal projection matrices of rank p [11, 12];

$$\text{Grass}(p, n) \simeq \left\{ X \in \mathbb{R}^{n \times n} \mid X^T = X, X^2 = X, \text{rank}(X) = p \right\} \quad (2.4)$$

$$= \left\{ X = YY^T \mid Y \in \text{St}(p, n) \right\}. \quad (2.5)$$

With this definition, the existing algorithms on the Grassmann manifold will be reformulated so as to be more feasible on the computer.

We proceed to the tangent space to the Grassmann manifold.

2.2 Tangent spaces

Proposition 2.1 *The tangent space $T_X \text{Grass}(p, n)$ at $X \in \text{Grass}(p, n)$ is expressed as*

$$T_X \text{Grass}(p, n) = \left\{ \xi \in \mathbb{R}^{n \times n} \mid \xi^T = \xi, \xi = \xi X + X\xi \right\}. \quad (2.6)$$

Further, let E_{ij} denote the $p \times (n - p)$ matrix whose (i, j) -component is 1 and the others 0. For $X \in \text{Grass}(p, n)$, let $Y \in \text{St}(p, n)$ and $Y_\perp \in \text{St}(n - p, n)$ be matrices such that $X = YY^T$ and $I_n - X = Y_\perp Y_\perp^T$, respectively. Then, the set of matrices $\xi_{ij} := \text{sym}(Y E_{ij} Y_\perp^T)$ with $i = 1, \dots, p$ and $j = 1, \dots, n - p$ forms a basis of the tangent space $T_X \text{Grass}(p, n)$, where $\text{sym}(\cdot)$ denotes the symmetric part of the matrix in the parentheses.

Proof Let ξ be an element of $T_X \text{Grass}(p, n)$. By differentiation it follows from $X^T = X$ and $X^2 = X$ that $\xi^T = \xi$ and $\xi = \xi X + X\xi$. Therefore,

$$T_X \text{Grass}(p, n) \subset \left\{ \xi \in \mathbb{R}^{n \times n} \mid \xi^T = \xi, \xi = \xi X + X\xi \right\}. \quad (2.7)$$

Let V_1 be the right-hand side of Eq. (2.7). We shall show that the dimension of V_1 is equal to that of $T_X \text{Grass}(p, n)$. Since any eigenvalue of an idempotent symmetric matrix is equal to 1 or 0, it follows from $\text{rank}(X) = p$ that there exists an $n \times n$ orthogonal matrix P such that

$$X = P\Lambda P^T, \quad (2.8)$$

where $\Lambda = \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}$. Let us put P in the form $P = (Y, Y_\perp)$, where Y and Y_\perp are $n \times p$ and $n \times (n - p)$ matrices, respectively. Then, Y and Y_\perp satisfy the conditions of this proposition. It follows from Eq. (2.8) that

$$V_1 = \left\{ \xi \in \mathbb{R}^{n \times n} \mid (P^T \xi P)^T = P^T \xi P, P^T \xi P = (P^T \xi P) \Lambda + \Lambda (P^T \xi P) \right\}. \quad (2.9)$$

Let V_2 denote the vector space $\text{Ad}_{P^{-1}}(V_1)$, where $\text{Ad}_{P^{-1}}(\xi) = P^{-1}\xi P$ for any $\xi \in V_1$. Since $\text{Ad}_{P^{-1}}(\xi) = P^T \xi P$, we obtain

$$V_2 = \left\{ \eta \in \mathbb{R}^{n \times n} \mid \eta^T = \eta, \eta = \eta \Lambda + \Lambda \eta \right\} = \left\{ \begin{pmatrix} 0 & D \\ D^T & 0 \end{pmatrix} \in \mathbb{R}^{n \times n} \mid D \in \mathbb{R}^{p \times (n-p)} \right\}. \quad (2.10)$$

Since the map $\text{Ad}_{P^{-1}}$ is an isomorphism from V_1 to V_2 , we have

$$\dim(V_1) = \dim(V_2) = p(n-p). \quad (2.11)$$

Thus, Eq. (2.6) is verified.

The set of matrices $\eta_{ij} := \text{sym} \begin{pmatrix} 0 & E_{ij} \\ 0 & 0 \end{pmatrix}$ with $i = 1, \dots, p$, $j = 1, \dots, n-p$ forms a basis of V_2 , so that the set of matrices $\text{Ad}_P(\eta_{ij})$ with $i = 1, \dots, p$, $j = 1, \dots, n-p$ forms a basis of V_1 . The $\text{Ad}_P(\eta_{ij})$ are written out to be

$$\text{Ad}_P(\eta_{ij}) = (Y, Y_\perp) \text{sym} \begin{pmatrix} 0 & E_{ij} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Y^T \\ Y_\perp^T \end{pmatrix} = \text{sym} (Y E_{ij} Y_\perp^T). \quad (2.12)$$

This completes the proof. \square

We have here to remark on the condition resulting from $\text{rank}(X) = p$. Since any eigenvalue of an idempotent symmetric matrix Z is equal to 1 or 0, it turns out that $\text{rank}(Z) = p$ if and only if $\text{tr}(Z) = p$. Then, the definition (2.4) of the Grassmann manifold is described also as

$$\text{Grass}(p, n) \simeq \left\{ X \in \mathbb{R}^{n \times n} \mid X^T = X, X^2 = X, \text{tr}(X) = p \right\}. \quad (2.13)$$

Though the condition $\text{tr}(\xi) = 0$ resulting from $\text{tr}(X) = p$ seems to be missing from the right-hand side of (2.6), it is a consequence of the condition $\xi = \xi X + X \xi$ with $X \in \text{Grass}(p, n)$. Indeed, we can obtain $X \xi X = 0$ from $\xi = \xi X + X \xi$, and thereby

$$\text{tr}(\xi) = \text{tr}(\xi X + X \xi) = \text{tr}(\xi X^2 + X^2 \xi) = 2 \text{tr}(X \xi X) = 0. \quad (2.14)$$

We here have to note that there is another description of the tangent space. For example, in Thm. 2.1 of [12], tangent vectors at $X \in \text{Grass}(p, n)$ are expressed, in terms of Lie brackets, as $[X, \eta]$ with $\eta \in \text{so}(n)$.

Since the Grassmann manifold is a submanifold of the matrix Euclidean space $\mathbb{R}^{n \times n}$, it is endowed with the Riemannian metric

$$\langle \xi, \eta \rangle_X := \text{tr}(\xi^T \eta) = \text{tr}(\xi \eta), \quad \xi, \eta \in T_X \text{Grass}(p, n), \quad (2.15)$$

which is induced from the natural metric on $\mathbb{R}^{n \times n}$,

$$\langle B, C \rangle := \text{tr}(B^T C), \quad B, C \in \mathbb{R}^{n \times n}. \quad (2.16)$$

Lemma 2.1 *The length of each basis matrix ξ_{ij} given in Prop. 2.1 is $1/\sqrt{2}$.*

Proof By the definition (2.15) of the Riemannian metric, the squared length of $\xi_{ij} = \text{sym}(Y E_{ij} Y_{\perp}^T)$ is given by and calculated as

$$\|\xi_{ij}\|^2 = \text{tr} \left(\left(\text{sym}(Y E_{ij} Y_{\perp}^T) \right)^2 \right) = \frac{1}{4} \left(\text{tr}(E_{ij} E_{ij}^T) + \text{tr}(E_{ij}^T E_{ij}) \right) = \frac{1}{2}. \quad (2.17)$$

This completes the proof. \square

By using Prop. 2.1 and Lemma 2.1, we can give the expression of the orthogonal projection onto the tangent space $T_X \text{Grass}(p, n)$.

Proposition 2.2 *The orthogonal projection operator π_{T_X} onto the tangent space $T_X \text{Grass}(p, n)$ at $X \in \text{Grass}(p, n)$ is given, for any $D \in \mathbb{R}^{n \times n}$, by*

$$\pi_{T_X}(D) = 2 \text{sym}(X \text{sym}(D) (I_n - X)). \quad (2.18)$$

Proof Let $\langle \cdot, \cdot \rangle$ denote the natural inner product (2.16) on $\mathbb{R}^{n \times n}$. By using the basis $\{\xi_{ij}\}$ in Prop. 2.1, the projection π_{T_X} is expressed and calculated, for any $D \in \mathbb{R}^{n \times n}$, as

$$\begin{aligned} \pi_{T_X}(D) &= \sum_{i,j} \left\langle D, \frac{\xi_{ij}}{\|\xi_{ij}\|} \right\rangle \frac{\xi_{ij}}{\|\xi_{ij}\|} \\ &= 2 \sum_{i,j} \text{tr} \left(D^T \text{sym}(Y E_{ij} Y_{\perp}^T) \right) \xi_{ij} \\ &= 2 \text{sym} \left(Y \left(\sum_{i,j} (Y^T \text{sym}(D) Y_{\perp})_{ij} E_{ij} \right) Y_{\perp}^T \right) \\ &= 2 \text{sym} \left(Y (Y^T \text{sym}(D) Y_{\perp}) Y_{\perp}^T \right) \\ &= 2 \text{sym}(X \text{sym}(D) (I_n - X)), \end{aligned} \quad (2.19)$$

where use has been made of the equality $Y Y^T + Y_{\perp} Y_{\perp}^T = I_n$ along with $X = Y Y^T$. \square

Another description of the projection operator is given in Prop. 2.1 of [12], which is put in the form of double brackets.

We proceed to the relation between tangent spaces to the Grassmann and to the Stiefel manifolds.

Lemma 2.2 *For a curve $X(t)$ on the Grassmann manifold, there exists a curve $Y(t)$ on the Stiefel manifold such that*

$$X(t) = Y(t)Y(t)^T, \quad Y(0)^T \dot{Y}(0) = 0. \quad (2.20)$$

Moreover, for this curve $Y(t)$, one has

$$\dot{Y}(0) = \dot{X}(0)Y(0). \quad (2.21)$$

Proof We note first that for a given $X(t)$ there exists a curve $Y_0(t) \in \text{St}(p, n)$ such that $X(t) = Y_0(t)Y_0(t)^T$ for all $t \in \mathbb{R}$. Let $Q(t)$ be any curve on the orthogonal group $O(p)$ which satisfies

$$\dot{Q}(0) = -Y_0(0)^T \dot{Y}_0(0) Q(0). \quad (2.22)$$

We note here that since the tangent space $T_Y \text{St}(p, n)$ is expressed as

$$T_Y \text{St}(p, n) = \left\{ \Delta \in \mathbb{R}^{n \times p} \mid \Delta^T Y + Y^T \Delta = 0 \right\}, \quad (2.23)$$

$Y_0(0)^T \dot{Y}_0(0)$ is skew-symmetric, so that $-Y_0(0)^T \dot{Y}_0(0) Q(0)$ is a tangent vector to $O(p)$ at $Q(0)$ (for the proof of Eq. (2.23), see [2, 6]). For such $Y_0(t)$ and $Q(t)$, we define a curve $Y(t)$ to be $Y(t) = Y_0(t)Q(t)$. Then we verify that

$$Y(t)Y(t)^T = Y_0(t)Q(t)Q(t)^T Y_0(t)^T = Y_0(t)Y_0(t)^T = X(t) \quad (2.24)$$

and that

$$Y(0)^T \dot{Y}(0) = Q(0)^T Y_0(0)^T \left(\dot{Y}_0(0) Q(0) + Y_0(0) \dot{Q}(0) \right) = 0. \quad (2.25)$$

Thus, Eq. (2.20) is proved. We proceed to prove Eq. (2.21). Differentiating $X(t) = Y(t)Y(t)^T$ with respect to t at $t = 0$, we obtain $\dot{X}(0) = \dot{Y}(0)Y(0)^T + Y(0)\dot{Y}(0)^T$. Multiplying this equation by $Y(0)$ from the right results in

$$\dot{X}(0)Y(0) = \dot{Y}(0). \quad (2.26)$$

This ends the proof. \square

We observe from (2.26) that for $\xi \in T_X \text{Grass}(p, n)$, the ξY is a tangent vector to $\text{St}(p, n)$ at Y with $X = YY^T$. This observation gives rise to the following proposition.

Proposition 2.3 *Let X be an element of $\text{Grass}(p, n)$ and ξ a tangent vector to $\text{Grass}(p, n)$ at X . Let Y be an element of $\text{St}(p, n)$ such that $X = YY^T$. Then, $\Delta := \xi Y$ is a tangent vector to $\text{St}(p, n)$ at Y such that $\xi = \Delta Y^T + Y \Delta^T$ and $Y^T \Delta = 0$.*

Proof From Eq. (2.6), we have $\xi = \xi X + X \xi = \xi Y Y^T + Y Y^T \xi$. Multiplying this equation by Y^T and Y from the left and the right, respectively, we obtain $Y^T \xi Y = 0$ and hence $Y^T \Delta = 0$. From Eq. (2.23), $\Delta = \xi Y$ proves to be a tangent vector to $\text{St}(p, n)$ at Y . Further, for $\Delta = \xi Y$, we have

$$\Delta Y^T + Y \Delta^T = \xi Y Y^T + Y Y^T \xi = \xi X + X \xi = \xi. \quad (2.27)$$

This completes the proof. \square

2.3 Geodesics

We shall find explicitly the exponential map on the Grassmann manifold by solving the geodesic equation.

Proposition 2.4 *The geodesic equation on the Grassmann manifold (2.4) is expressed as*

$$\ddot{X} + 2\dot{X}^2 - 4\dot{X}X\dot{X} = 0. \quad (2.28)$$

We note also that Eq. (2.28) turns out to be equivalent to an equation given in [12] (see (5.1) in Section 5). For the proof of (2.28), see Appendix A.

We can describe solutions to the geodesic equation (2.28) by means of Lemma 2.2.

Proposition 2.5 *Let $X(t)$ be a geodesic on the Grassmann manifold and $Y(t)$ a curve on the Stiefel manifold such that*

$$X(t) = Y(t)Y(t)^T, \quad Y(0)^T \dot{Y}(0) = 0, \quad (2.29)$$

where the existence of such a $Y(t)$ is proved by Lemma 2.2, and where $\dot{X}(0)$ and $\dot{Y}(0)$ are related by (2.21). Let U , Σ , and V be the factors of the thin singular value decomposition $[10, 17]$ of $\dot{Y}(0)$, that is,

$$\dot{Y}(0) = U\Sigma V^T, \quad (2.30)$$

where $U \in \text{St}(p, n)$, $V \in O(p)$, and where Σ is a $p \times p$ diagonal matrix whose diagonal entries are nonnegative. Then, $X(t)$ is expressed as

$$X(t) = \frac{1}{2} (YV U) \begin{pmatrix} I_p + \cos 2\Sigma t & \sin 2\Sigma t \\ \sin 2\Sigma t & I_p - \cos 2\Sigma t \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix}, \quad (2.31)$$

where Y is used instead of $Y(0)$ for short.

Proof Let $X_1(t)$ denote the right-hand side of (2.31). Differentiating $X_1(t)$ with respect to t , we obtain

$$\dot{X}_1(t) = (YV U) \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} \begin{pmatrix} -\sin 2\Sigma t & \cos 2\Sigma t \\ \cos 2\Sigma t & \sin 2\Sigma t \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix} \quad (2.32)$$

and

$$\ddot{X}_1(t) = 2(YV U) \begin{pmatrix} \Sigma^2 & 0 \\ 0 & \Sigma^2 \end{pmatrix} \begin{pmatrix} -\cos 2\Sigma t & -\sin 2\Sigma t \\ -\sin 2\Sigma t & \cos 2\Sigma t \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix}. \quad (2.33)$$

Then, a straightforward calculation shows that $X_1(t)$ satisfies the geodesic equation (2.28). For initial values of $X_1(t)$ and $\dot{X}_1(t)$, we have

$$X_1(0) = YV V^T Y^T = Y Y^T = X(0) \quad (2.34)$$

and

$$\dot{X}_1(0) = U\Sigma V^T Y^T + YV\Sigma U^T = \dot{Y}(0)Y^T + Y\dot{Y}(0)^T = \dot{X}(0). \quad (2.35)$$

Thus, the theorem on existence and uniqueness of solutions to ordinary differential equations ensures that $X(t) = X_1(t)$. \square

We note that the solution (2.31) is the deduction of Thm. 2.3 in [6], which provides the solution of the geodesic equation on $\text{Grass}(p, n)$ viewed as $\text{St}(p, n)/O(p)$.

2.4 Retraction

Here we introduce the notion of a retraction after [2, 4], which provides a way to determine a next iterate with a given search direction.

Definition 2.1 *Let M and TM be a manifold and the tangent bundle of M , respectively. Let $R : TM \rightarrow M$ be a smooth map and R_x the restriction of R to $T_x M$. The R is called a retraction on M , if it has the following properties:*

1. $R_x(0_x) = x$, where 0_x denotes the zero element of $T_x M$.
2. With the canonical identification $T_{0_x} T_x M \simeq T_x M$, R_x satisfies

$$DR_x(0_x) = \text{id}_{T_x M}, \quad (2.36)$$

where $DR_x(0_x)$ denotes the derivative of R_x at 0_x , and $\text{id}_{T_x M}$ the identity map on $T_x M$.

A typical example of a retraction on a Riemannian manifold M is the exponential map on M . From Lemma 2.2 and Prop. 2.5, we can put the exponential map on the Grassmann manifold in the form

$$\text{Exp}_X(\xi) = \frac{1}{2} (YV U) \begin{pmatrix} I_p + \cos 2\Sigma & \sin 2\Sigma \\ \sin 2\Sigma & I_p - \cos 2\Sigma \end{pmatrix} \begin{pmatrix} V^T Y^T \\ U^T \end{pmatrix}, \quad (2.37)$$

where $\xi \in T_X \text{Grass}(p, n)$, $X = YY^T$, and where U, Σ , and V are the factors of the thin singular value decomposition of ξY ; $\xi Y = U \Sigma V^T$. We call the map $R : T\text{Grass}(p, n) \rightarrow \text{Grass}(p, n)$, determined by $R_X = \text{Exp}_X$, the exponential retraction.

There is another retraction on the Grassmann manifold, which is based on the QR decomposition. The QR decomposition of a full-rank $n \times p$ matrix B is put in the form [10, 17]

$$B = QR, \quad Q \in \text{St}(p, n), \quad R \in S_{\text{upp}}^+(p), \quad (2.38)$$

where $S_{\text{upp}}^+(p)$ denotes the set of all $p \times p$ upper triangular matrices with strictly positive diagonal entries. Let $\text{qf}(B)$ denote the Q factor of the QR decomposition of $B = QR$, that is, $\text{qf}(B) = Q$.

Before giving a retraction based on the QR decomposition, we show the following lemmas.

Lemma 2.3 *Let Y_1 and Y_2 be elements of $\text{St}(p, n)$. If there exists a $p \times p$ matrix Q such that $Y_1 = Y_2 Q$, then Q is an orthogonal matrix.*

Proof Since $Y_1^T Y_1 = Y_2^T Y_2 = I_p$, we obtain

$$I_p = Y_1^T Y_1 = Q^T Y_2^T Y_2 Q = Q^T Q. \quad (2.39)$$

This ends the proof. \square

Lemma 2.4 *Let B and Q be elements of $\mathbb{R}^{n \times p}$ and of $O(p)$, respectively. Then, one has*

$$\text{qf}(B) (\text{qf}(B))^T = \text{qf}(BQ) (\text{qf}(BQ))^T. \quad (2.40)$$

Proof Let $B = Q_1 R_1$ and $BQ = Q_2 R_2$ be the QR decompositions of B and BQ , respectively. Note that $Q_1, Q_2 \in \text{St}(p, n)$ and $R_1, R_2 \in \text{GL}(p)$. What we should show is that $Q_1 Q_1^T = Q_2 Q_2^T$. From these QR decompositions, we obtain

$$Q_2 = BQ R_2^{-1} = Q_1 (R_1 Q R_2^{-1}). \quad (2.41)$$

It follows from Lemma 2.3 that $R_1 Q R_2^{-1}$ is an orthogonal matrix, so that

$$Q_2 Q_2^T = Q_1 (R_1 Q R_2^{-1}) (R_1 Q R_2^{-1})^T Q_1^T = Q_1 Q_1^T. \quad (2.42)$$

This completes the proof. \square

We give the retraction based on the QR decomposition as follows:

Proposition 2.6 *Let R_X be a map of $T_X \text{Grass}(p, n)$ to $\text{Grass}(p, n)$ defined by*

$$R_X(\xi) = \text{qf}((I_n + \xi)Y) (\text{qf}((I_n + \xi)Y))^T, \quad \xi \in T_X \text{Grass}(p, n), \quad (2.43)$$

where $Y \in \text{St}(p, n)$ satisfies $X = YY^T$. Then, the collection of R_X for all $X \in \text{Grass}(p, n)$ forms a retraction $R : T\text{Grass}(p, n) \rightarrow \text{Grass}(p, n)$.

Proof We first show that the right-hand side of (2.43) is independent of the choice of Y . Let Y_1 and Y_2 be elements of $\text{St}(p, n)$, satisfying $X = Y_1 Y_1^T = Y_2 Y_2^T$. Since $\text{Grass}(p, n) \simeq \text{St}(p, n)/O(p)$, Y_1 and Y_2 are related by $Y_2 = Y_1 Q$ with $Q \in O(p)$. Since $(I_n + \xi)Y_1$ is of full-rank, Eq. (2.40) with $B = (I_n + \xi)Y_1$ together with $Y_2 = Y_1 Q$ implies that

$$\text{qf}((I_n + \xi)Y_1) (\text{qf}((I_n + \xi)Y_1))^T = \text{qf}((I_n + \xi)Y_2) (\text{qf}((I_n + \xi)Y_2))^T. \quad (2.44)$$

Thus, the right-hand side of (2.43) is well-defined.

We move to show that R is a retraction. We first prove that $R_X(\xi) \in \text{Grass}(p, n)$. A straightforward calculation shows that $R_X(\xi)^2 = R_X(\xi)$ and $R_X(\xi)^T = R_X(\xi)$. Further, it is straightforward to show that

$$\text{tr}(R_X(\xi)) = \text{tr}((\text{qf}((I_n + \xi)Y))^T \text{qf}((I_n + \xi)Y)) = \text{tr}(I_p) = p. \quad (2.45)$$

The remaining task is to show that the R_X given by (2.43) satisfies the two conditions imposed in Definition 2.1. The first condition in Definition 2.1 is easy to verify;

$$R_X(0) = \text{qf}(Y) (\text{qf}(Y))^T = YY^T = X. \quad (2.46)$$

To verify the second condition, we take the derivative of R_X . Since $\xi Y \in T_Y \text{St}(p, n)$ and $D \text{qf}(Y)[\Delta] = \Delta$ for any $\Delta \in T_Y \text{St}(p, n)$ [2], the derivative of R_X at 0 with respect to ξ is given by and written out as

$$\begin{aligned} DR_X(0)[\xi] &= (D \text{qf}(Y)[\xi Y]) (\text{qf}(Y))^T + \text{qf}(Y) (D \text{qf}(Y)[\xi Y])^T \\ &= \xi Y Y^T + Y (\xi Y)^T = \xi X + X \xi = \xi. \end{aligned} \quad (2.47)$$

Thus the second condition in Definition 2.1 is confirmed for the present R_X . This completes the proof. \square

We call the R defined through (2.43) the QR-based retraction.

We here note that another form of the QR-based retraction is given in (2.69) of [12], which is defined by

$$\mu_X^{\text{QR}}(\xi) = \text{qf}(I + [\xi, X])X (\text{qf}(I + [\xi, X]))^T, \quad (2.48)$$

where $[\cdot, \cdot]$ denotes the Lie bracket. At the cost of computing $Y \in \text{St}(p, n)$ such that $X = YY^T$, our QR-based retraction (2.43) needs only the QR decomposition of the $n \times p$ matrix $(I + \xi)Y$. In contrast with this, the retraction (2.48) needs the QR decomposition of the $n \times n$ matrix $I + [\xi, X]$. An algorithm to compute $Y \in \text{St}(p, n)$ satisfying $X = YY^T$ is provided as Algorithm 3.2 in the next section.

2.5 The gradient and the Hessian of a function

Let $F : \text{Grass}(p, n) \rightarrow \mathbb{R}$ be a smooth function. The gradient and the Hessian of F are inevitable in optimization methods. The gradient, $\text{grad } F(X)$, of the function F at $X \in \text{Grass}(p, n)$ is defined to be a unique tangent vector which satisfies

$$\langle \text{grad } F(X), \xi \rangle_X = DF(X)[\xi] \quad (2.49)$$

for any $\xi \in T_X \text{Grass}(p, n)$.

Proposition 2.7 *The gradient of F at $X \in \text{Grass}(p, n)$ is expressed as*

$$\text{grad } F(X) = F_X X + X F_X - 2X F_X X, \quad (2.50)$$

where F_X denotes the $n \times n$ matrix whose (i, j) component is $\partial F(X)/\partial X_{ij}$.

Proof Since $\text{Grass}(p, n)$ is a Riemannian submanifold of $\mathbb{R}^{n \times n}$ endowed with the induced metric, $\text{grad } F(X)$ is equal to the projection of the Euclidean gradient of F at X onto $T_X \text{Grass}(p, n)$. Hence, by using the projection π_{T_X} given in (2.18), we obtain

$$\text{grad } F(X) = \pi_{T_X}(F_X) = 2 \text{sym}(X F_X (I_n - X)) = F_X X + X F_X - 2X F_X X. \quad (2.51)$$

This completes the proof. \square

We proceed to the Hessian of a function F at $X \in \text{Grass}(p, n)$. The Hessian of F at X is defined to be a symmetric linear map of the tangent space $T_X \text{Grass}(p, n)$ through

$$\langle \text{Hess } F(X)[\xi], \xi \rangle_X = \left. \frac{D}{dt} \frac{dF}{dt}(X(t)) \right|_{t=0}, \quad (2.52)$$

where $X(t)$ is a smooth curve passing X at $t = 0$ with $\dot{X}(0) = \xi$ and where $\frac{D}{dt}$ denotes the covariant derivation of a vector field along a curve. If $X(t)$ is chosen as a geodesic passing X at $t = 0$, the defining equation (2.52) takes the form

$$\langle \text{Hess } F(X)[\xi], \xi \rangle_X = \left. \frac{d^2}{dt^2} F(X(t)) \right|_{t=0}. \quad (2.53)$$

Proposition 2.8 *Let ξ be a tangent vector at $X \in \text{Grass}(p, n)$ and let $F_{XX}[\xi]$ denote the $n \times n$ matrix whose (i, j) component is $\sum_{k,l=1}^n (\partial^2 F(X)/\partial X_{ij} \partial X_{kl}) \xi_{kl}$. The Hessian of F at X acts on ξ as a linear map on $T_X \text{Grass}(p, n)$ by*

$$\text{Hess } F(X)[\xi] = 2 \text{sym} \left(X \text{sym}(F_{XX}[\xi] + 4F_X \xi X - 2F_X \xi) (I_n - X) \right). \quad (2.54)$$

Proof Let $X(t)$ be the geodesic emanating from $X(0) = X \in \text{Grass}(p, n)$ in the direction of $\dot{X}(0) = \xi \in T_X \text{Grass}(p, n)$. Note that $X(t)$ satisfies the geodesic equation (2.28) and hence

$$\ddot{X}(0) + 2\dot{X}(0)^2 - 4\dot{X}(0)X(0)\dot{X}(0) = 0. \quad (2.55)$$

By computing the right-hand side of (2.53) along with (2.55), we find that $\langle \text{Hess } F(X)[\xi], \xi \rangle_X$ is written out as

$$\begin{aligned} \langle \text{Hess } F(X)[\xi], \xi \rangle_X &= \sum_{i,j,k,l} \frac{\partial^2 F}{\partial X_{ij} \partial X_{kl}}(X) \dot{X}(0)_{ij} \dot{X}(0)_{kl} + \text{tr} \left(F_X \ddot{X}(0) \right) \\ &= \text{tr} (F_{XX}[\xi] \xi) + 2 \text{tr} \left(-F_X \xi^2 + 2F_X \xi X \xi \right). \end{aligned} \quad (2.56)$$

Since the Hessian operator is symmetric and linear, for any $\xi, \eta \in T_X \text{Grass}(p, n)$, the symmetric form in ξ and η is expressed and written out as

$$\begin{aligned} &\langle \text{Hess } F(X)[\xi], \eta \rangle_X \\ &= \frac{1}{2} \left(\langle \text{Hess } F(X)[\xi + \eta], \xi + \eta \rangle_X - \langle \text{Hess } F(X)[\xi], \xi \rangle_X - \langle \text{Hess } F(X)[\eta], \eta \rangle_X \right) \\ &= \text{tr} \left((F_{XX}[\xi] + 4F_X \xi X - 2F_X \xi) \eta \right) \\ &= \langle \pi_{T_X} (F_{XX}[\xi] + 4F_X \xi X - 2F_X \xi), \eta \rangle_X. \end{aligned} \quad (2.57)$$

Hence, we have

$$\begin{aligned} \text{Hess } F(X)[\xi] &= \pi_{T_X} (F_{XX}[\xi] + 4F_X \xi X - 2F_X \xi) \\ &= 2 \text{sym} \left(X \text{sym}(F_{XX}[\xi] + 4F_X \xi X - 2F_X \xi) (I_n - X) \right). \end{aligned} \quad (2.58)$$

This completes the proof. \square

We here remark that the gradient and the Hessian of a function at $X \in \text{Grass}(p, n)$ are given in Thm. 2.4 of [12] in terms of double brackets. The difference between ours and theirs in the description of the gradient and the Hessian results from the difference in the description of tangent vectors.

3 Optimization algorithms on $\text{Grass}(p, n)$

This section deals with the steepest descent method and Newton's method on $\text{Grass}(p, n)$ on the basis of the geometric setting up given in Section 2. As in Section 2, some results in this section are related to [12]. We again make remarks on the relation to theirs if necessary.

3.1 The steepest descent and Newton's methods on a general Riemannian manifold M

Before specializing in optimization algorithms on $\text{Grass}(p, n)$, we make a brief review of the problem of minimizing the objective function f defined on a Riemannian manifold M , following [2]. The steepest descent and Newton's methods on M have the following common framework.

Algorithm 3.1 The framework of the steepest descent and Newton's methods on a Riemannian manifold M

- 1: Choose an initial point $x_0 \in M$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: Compute the search direction $\eta_k \in T_{x_k}M$ and the step size $t_k > 0$.
 - 4: Compute the next iterate $x_{k+1} = R_{x_k}(t_k \eta_k)$, where R is a retraction on M .
 - 5: **end for**
-

The choice of the search direction $\eta_k \in T_{x_k}M$ and the step size t_k in Step 3 of Algorithm 3.1 characterizes the steepest descent and Newton's methods. In the steepest descent method, η_k is computed as

$$\eta_k = -\text{grad } f(x_k), \quad (3.1)$$

while in Newton's method, η_k is determined as the solution of Newton's equation

$$\text{Hess } f(x_k)[\eta_k] = -\text{grad } f(x_k), \quad (3.2)$$

where the Hessian $\text{Hess } f(x)$ of f at x is in general related with the covariant derivative $\nabla_\eta \text{grad } f$ for an affine connection ∇ by

$$\text{Hess } f(x)[\eta] := \nabla_\eta \text{grad } f. \quad (3.3)$$

We note that Eq. (2.52) is a consequence of (3.3). For the step size t_k , we adopt the Armijo step size in the steepest descent method, which is determined, for given parameters $\bar{\alpha} > 0$, β , $\sigma \in (0, 1)$, by $t_k := \beta^m \bar{\alpha}$ in such a way that m may be the smallest nonnegative integer satisfying

$$f(x) - f(R_x(\beta^m \bar{\alpha} \eta)) \geq -\sigma \langle \text{grad } f(x), \beta^m \bar{\alpha} \eta \rangle_x, \quad (3.4)$$

where $\langle \cdot, \cdot \rangle_x$ denotes the inner product on $T_x M$. In Newton's method, we fix $t_k := 1$ for any k .

According to [2], convergence results for the steepest descent and Newton's methods are, respectively, stated as follows:

Theorem 3.1 *Let $\{x_k\}$ be an infinite sequence of iterates generated by the steepest descent method on the Riemannian manifold M with the objective function f . If M is compact, then*

$$\lim_{k \rightarrow \infty} \|\text{grad } f(x_k)\|_{x_k} = 0. \quad (3.5)$$

Theorem 3.2 *Let $x_c \in M$ be a critical point of f ; $\text{grad } f(x_c) = 0$. Assume that $\text{Hess } f(x_c)$ is non-degenerate at $x_c \in M$. Then there exists a neighborhood U of x_c in M such that for all $x_0 \in U$ the sequence $\{x_k\}$ generated by Newton's method converges quadratically to x_c .*

3.2 An algorithmic setting up on $\text{Grass}(p, n)$

In what follows, we specialize in $\text{Grass}(p, n)$. Let F be a smooth objective function on $\text{Grass}(p, n)$. So far we have obtained all geometric requisites for the steepest descent method on the Grassmann manifold. However, from the viewpoint of algorithm, we have to give a procedure to construct $Y_k \in \text{St}(p, n)$ such that $X_k = Y_k Y_k^T$ for each $X_k \in \text{Grass}(p, n)$. To this end, we prove the following proposition.

Proposition 3.1 *Let $X = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ be an element of $\text{Grass}(p, n)$, where \mathbf{x}_k are column vectors of X . Assume that the set of p vectors $\{\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_p}\}$ is linearly independent, where $1 \leq i_1 < \dots < i_p \leq n$, and denote the $n \times p$ matrix $(\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_p})$ by Z_0 . Let Q and R be the factors of the QR decomposition of Z_0 :*

$$Z_0 = QR, \quad Q \in \text{St}(p, n), \quad R \in S_{\text{upp}}^+(p). \quad (3.6)$$

Then, X is expressed as

$$X = QQ^T. \quad (3.7)$$

Proof By permuting the columns of X , we obtain the matrix $(\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_p}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_{n-p}})$, where the set $\{j_1, \dots, j_{n-p}\}$ of indices is the complement of the subset $\{i_1, \dots, i_p\}$ in $\{1, \dots, n\}$ with $j_1 < \dots < j_{n-p}$. Since a transposition of the column vectors, and hence the product of such transpositions, can be represented as an orthogonal matrix, there exists an $n \times n$ orthogonal matrix P such that $XP = (\mathbf{x}_{i_1}, \dots, \mathbf{x}_{i_p}, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_{n-p}})$.

We partition the matrix XP into

$$XP = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix}, \quad (3.8)$$

where $X_{11} \in \mathbb{R}^{p \times p}$, $X_{12} \in \mathbb{R}^{p \times (n-p)}$, $X_{21} \in \mathbb{R}^{(n-p) \times p}$, $X_{22} \in \mathbb{R}^{(n-p) \times (n-p)}$. Then, X is brought into

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} P^T. \quad (3.9)$$

Since $X^T = X$ and $X^2 = X$, we obtain $X^T X = X$, which is written out as

$$P \begin{pmatrix} X_{11}^T X_{11} + X_{21}^T X_{21} & X_{11}^T X_{12} + X_{21}^T X_{22} \\ X_{12}^T X_{11} + X_{22}^T X_{21} & X_{12}^T X_{12} + X_{22}^T X_{22} \end{pmatrix} P^T = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} P^T. \quad (3.10)$$

Multiplying Eq. (3.10) by P^T and $P \begin{pmatrix} I_p \\ 0 \end{pmatrix}$ from the left and the right, respectively, we obtain

$$\begin{pmatrix} X_{11}^T X_{11} + X_{21}^T X_{21} \\ X_{12}^T X_{11} + X_{22}^T X_{21} \end{pmatrix} = P^T \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}. \quad (3.11)$$

Since $\text{rank}(XP) = \text{rank} \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} = p$, there exists a $p \times (n-p)$ matrix C such that

$$\begin{pmatrix} X_{12} \\ X_{22} \end{pmatrix} = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} C, \quad (3.12)$$

which means that each of \mathbf{x}_{j_k} with $1 \leq k \leq n - p$ is a linear combination of \mathbf{x}_{i_l} with l ranging from 1 to p . Eq. (3.6) is then written as

$$Z_0 = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} = QR. \quad (3.13)$$

Since R is invertible, we obtain from (3.13)

$$Q = \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} R^{-1}. \quad (3.14)$$

By using Eqs. (3.11)–(3.14) and (3.9), the QQ^T is expressed and calculated as

$$\begin{aligned} QQ^T &= \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} R^{-1} (R^{-1})^T \begin{pmatrix} X_{11}^T & X_{21}^T \end{pmatrix} \\ &= \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} ((QR)^T (QR))^{-1} \left(P^T \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} \right)^T P^T \\ &= \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} (X_{11}^T X_{11} + X_{21}^T X_{21})^{-1} (X_{11}^T X_{11} + X_{21}^T X_{21} X_{11}^T X_{11} C + X_{21}^T X_{21} C) P^T \\ &= \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} (I_p \ C) P^T = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} P^T = X. \end{aligned} \quad (3.15)$$

This completes the proof. \square

In practice, however, since the leftmost p columns of X are often linearly independent, the following corollary is of great use.

Corollary 3.1 *Let X be an element of $\text{Grass}(p, n)$. Assume that the p leftmost columns of X are linearly independent. Let Q and R be the factors of the QR decomposition of the $n \times p$ matrix whose columns are the p leftmost columns of X :*

$$X \begin{pmatrix} I_p \\ 0 \end{pmatrix} = QR, \quad Q \in \text{St}(p, n), \ R \in S_{\text{upp}}^+(p). \quad (3.16)$$

Then, we have

$$X = QQ^T. \quad (3.17)$$

According to Prop. 3.1, for given $X \in \text{Grass}(p, n)$ and $\xi \in T_X \text{Grass}(p, n)$, we can obtain $Y \in \text{St}(p, n)$ and $\Delta \in T_Y \text{St}(p, n)$ satisfying $X = YY^T$ and $\xi = \Delta Y^T + Y \Delta^T$ (see Prop. 2.3), the algorithm for which is stated as follows:

Algorithm 3.2 Method for computing suitable $Y \in \text{St}(p, n)$ and $\Delta \in T_Y \text{St}(p, n)$ for given $X \in \text{Grass}(p, n)$ and $\xi \in T_X \text{Grass}(p, n)$

- 1: Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ denote columns of X from the left, that is, $X = (\mathbf{x}_1, \dots, \mathbf{x}_n)$.
 - 2: Set $i = 0$ and $j = 1$.
 - 3: **while** $i < p$ **do**
 - 4: **if** $\{\mathbf{x}'_1, \dots, \mathbf{x}'_i, \mathbf{x}_j\}$ is linearly independent **then**
 - 5: $\mathbf{x}'_{i+1} = \mathbf{x}_j$ and $i = i + 1$.
 - 6: **end if**
 - 7: $j = j + 1$.
 - 8: **end while**
 - 9: Set $Z_0 = (\mathbf{x}'_1, \dots, \mathbf{x}'_p)$.
 - 10: Compute $Y = \text{qf}(Z_0)$.
 - 11: Compute $\Delta = \xi Y$.
-

The case in which the algorithm does not numerically work well will be described in Section 4.1. We can find a way to avoid such an inconvenience.

3.3 The steepest descent and Newton's methods on $\text{Grass}(p, n)$

Taking in Algorithm 3.2, we obtain the algorithm for the steepest descent method for the problem of minimizing the objective function F on the Grassmann manifold as follows:

Algorithm 3.3 Steepest descent method on the Grassmann manifold $\text{Grass}(p, n)$

- 1: Choose an initial point $X_0 \in \text{Grass}(p, n)$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: Compute the search direction $\eta_k = -(F_{X_k} X_k + X_k F_{X_k} - 2X_k F_{X_k} X_k)$ and the Armijo step size $t_k > 0$.
 - 4: Compute $Y_k \in \text{St}(p, n)$ and $\Delta_k \in T_{Y_k} \text{St}(p, n)$ by using Algorithm 3.2.
 - 5: Compute the next iterate $X_{k+1} = R_{X_k}(t_k \eta_k)$, where R is a retraction on $\text{Grass}(p, n)$.
 - 6: **end for**
-

In the above algorithm, as is seen already, a possible choice of the retraction R is the exponential retraction (2.37) or the QR-based retraction (2.43). If we choose (2.37), we need both Y_k and Δ_k . In contrast with this, if we adopt (2.43), we use Y_k only.

Since the Grassmann manifold $\text{Grass}(p, n)$ is compact, Algorithm 3.3 applied to a problem on $\text{Grass}(p, n)$ generates a sequence $\{X_k\}$ converging to a critical point of F because of Thm. 3.1.

Now we describe Newton's method on the Grassmann manifold. Let F be a smooth objective function on $\text{Grass}(p, n)$. From Eqs. (2.50) and (2.54), Newton's equation (3.2) on the Grassmann manifold is expressed as

$$\begin{aligned} & 2 \text{sym}(X_k \text{sym}(F_{X_k X_k}[\eta_k] + 4F_{X_k} \eta_k X_k - 2F_{X_k} \eta_k)(I_n - X_k)) \\ & = -(F_{X_k} X_k + X_k F_{X_k} - 2X_k F_{X_k} X_k), \end{aligned} \quad (3.18)$$

where $F_{X_k X_k}[\eta_k]$ denotes the $n \times n$ symmetric matrix $F_{XX}[\eta]$ evaluated at $X = X_k$ for $\eta = \eta_k$. Newton's method on the Grassmann manifold is expressed as follows:

Algorithm 3.4 Newton's method on the Grassmann manifold $\text{Grass}(p, n)$

- 1: Choose an initial point $X_0 \in \text{Grass}(p, n)$.
- 2: **for** $k = 0, 1, 2, \dots$ **do**
- 3: Solve Newton's equation

$$\begin{aligned} & 2 \text{sym}(X_k \text{sym}(F_{X_k X_k}[\eta_k] + 4F_{X_k} \eta_k X_k - 2F_{X_k} \eta_k)(I_n - X_k)) \\ & = -(F_{X_k} X_k + X_k F_{X_k} - 2X_k F_{X_k} X_k) \end{aligned} \quad (3.19)$$

for the unknown $\eta_k \in T_{X_k} M$, where $F_{X_k X_k}[\eta_k]$ is the $n \times n$ matrix $F_{XX}[\eta]$ evaluated at $X = X_k$ for $\eta = \eta_k$.

- 4: Compute $Y_k \in \text{St}(p, n)$ and $\Delta_k \in T_{Y_k} \text{St}(p, n)$ by using Algorithm 3.2.
 - 5: Compute the next iterate $X_{k+1} = R_{X_k}(\eta_k)$, where R is a retraction on $\text{Grass}(p, n)$.
 - 6: **end for**
-

3.4 Application to the Rayleigh quotient cost function

In what follows, we apply the steepest descent method to the Rayleigh quotient cost function $F(Y) = \text{tr}(Y^T A Y) / 2$, where A is an $n \times n$ symmetric matrix and $Y \in \mathbb{R}^{n \times p}$ is subject to the condition that $Y^T Y = I_p$. On account of the constraints on Y , the function F can be viewed as defined on the Stiefel manifold. Since the F has $O(p)$ invariance, we can bring the problem into that on the Grassmann manifold by putting $X = Y Y^T$:

Problem 3.1

$$\text{minimize } F(X) := \frac{1}{2} \text{tr}(AX), \quad (3.20)$$

$$\text{subject to } X \in \text{Grass}(p, n). \quad (3.21)$$

To apply the steepest descent method on the Grassmann manifold to Problem 3.1, we need the gradient of $F(X)$, which is easily found from Prop. 2.7 to take the form

$$\text{grad } F(X) = \text{sym}(AX) - XAX. \quad (3.22)$$

The algorithm of the steepest descent method for Problem 3.1 is stated as follows:

Algorithm 3.5 Steepest descent method for Problem 3.1

- 1: Choose an initial point $X_0 \in \text{Grass}(p, n)$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: Compute the search direction $\eta_k = -(\text{sym}(AX_k) - X_k A X_k)$ and the Armijo step size $t_k > 0$.
 - 4: Compute $Y_k \in \text{St}(p, n)$ by using Step 1-10 in Algorithm 3.2.
 - 5: Compute the next iterate $X_{k+1} = R_{X_k}(t_k \eta_k)$, where R is the QR-based retraction.
 - 6: **end for**
-

We note that for the current point $X_k \in \text{Grass}(p, n)$ and the search direction $\xi_k \in T_{X_k} \text{Grass}(p, n)$, the QR-based retraction does not need $\Delta_k \in T_{Y_k} \text{St}(p, n)$, which is referred to in Step 11 of Algorithm 3.2.

We apply Newton's method to Problem 3.1. Since the objective function is simple, the associated Newton's equation is put in a simple form.

Proposition 3.2 *Newton's equation for Problem 3.1 is written out as*

$$\eta B + B \eta = C, \quad (3.23)$$

where $B := A - AX - XA$, $C := 2(-\text{sym}(AX) + XAX)$.

Proof For the objective function $F(X) := \text{tr}(AX)/2$ on $\text{Grass}(p, n)$, the quantities we need for Newton's equation are $F_X = A/2$ and $\partial^2 F(X)/\partial X_{ij} \partial X_{kl} = 0$, $i, j, k, l = 1, \dots, n$. Then, it follows from Eq. (2.54) that

$$\begin{aligned} \text{Hess } F(X)[\eta] &= 2 \text{sym}(X \text{sym}(2A\eta X - A\eta)(I_n - X)) \\ &= \text{sym}(X(\eta A - A\eta)(I_n - X)) \\ &= \text{sym}(\eta(A - AX - XA)). \end{aligned} \quad (3.24)$$

Substituting Eqs. (3.22) and (3.24) into Eq. (3.2) results in Eq. (3.23). This completes the proof. \square

We point out that Eq. (3.23) is known as a Lyapunov equation for η [9]. Since an efficient algorithm is known for solving the Lyapunov equation, Newton's method applied to $F(X) = \text{tr}(AX)/2$ works effectively.

Newton's method applied to Problem 3.1 is stated in Algorithm 3.6.

Algorithm 3.6 Newton's method for Problem 3.1

- 1: Choose an initial point $X_0 \in \text{Grass}(p, n)$.
 - 2: **for** $k = 0, 1, 2, \dots$ **do**
 - 3: Solve Newton's equation

$$\eta_k B_k + B_k \eta_k = C_k \quad (3.25)$$
 for the unknown $\eta_k \in T_{X_k} M$, where $B_k := A - AX_k - X_k A$, $C_k := 2(-\text{sym}(AX_k) + X_k AX_k)$.
 - 4: Compute $Y_k \in \text{St}(p, n)$ by using Step 1-10 in Algorithm 3.2.
 - 5: Compute the next iterate $X_{k+1} = R_{X_k}(\eta_k)$, where R is the QR-based retraction.
 - 6: **end for**
-

As in Algorithm 3.5, $\Delta_k \in T_{Y_k} \text{St}(p, n)$ is not needed in performing Algorithm 3.6. Further, we have to note that Eq. (3.25) in Algorithm 3.6 does not always have a unique solution. In fact, if A has a degenerate eigenvalue, we encounter such a case. In practice, however, this does not matter in performing the algorithm. We will discuss these issues in detail later in Subsections 3.6 and 3.7.

We compare our algorithms with those given in [12] for the Rayleigh quotient cost function. In their algorithms, they compute $\Theta_k \in SO(n)$ such that

$$X_k = \Theta_k^T \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} \Theta_k. \quad (3.26)$$

The Θ corresponds to $(Y, Y_\perp)^T$ in our notation. Our algorithm needs to compute Y only, a part of Θ .

3.5 Global optimal solution for the Rayleigh quotient cost function

To characterize a global optimal solution to Problem 3.1, we now study the properties of the function $F(X) = \text{tr}(AX)/2$ at a critical point. Let the eigenvalues of A be arranged in increasing order, $\lambda_1 \leq \dots \leq \lambda_n$.

Proposition 3.3 *A point $X \in \text{Grass}(p, n)$ is a critical point of $F(X) = \text{tr}(AX)/2$, if and only if X is expressed as*

$$X = YY^T, \quad Y \in \text{St}(p, n), \quad (3.27)$$

where each column of Y is a normalized eigenvector of A .

Proof To start with, we note that $\text{grad } F(X) = \text{sym}(AX) - XAX$ from Eq. (3.22). If $X \in \text{Grass}(p, n)$ is a critical point of F , then

$$\text{sym}(AX) - XAX = 0. \quad (3.28)$$

Multiplying Eq. (3.28) by X from the right yields

$$AX = XAX. \quad (3.29)$$

Since XAX is symmetric, Eq. (3.29) implies that AX is symmetric as well, so that

$$AX = XA, \quad (3.30)$$

that is, A and X commute. It then follows that they are simultaneously diagonalizable, that is, there exists an $n \times n$ orthogonal matrix P such that

$$P^T AP = \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_n}) \quad (3.31)$$

and

$$P^T XP = \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}, \quad (3.32)$$

where $i_1, \dots, i_n \in \{1, \dots, n\}$ are distinct integers. Multiplying Eq. (3.31) and (3.32) by P from the left results in

$$AP = P \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_n}) \quad (3.33)$$

and

$$XP = P \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}, \quad (3.34)$$

respectively. We partition the matrix P into

$$P = (Y, Y_\perp), \quad Y \in \text{St}(p, n), \quad Y_\perp \in \text{St}(p, n-p) \quad (3.35)$$

to obtain, from (3.33) and (3.34),

$$AY = Y \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p}) \quad (3.36)$$

and

$$XY = Y, \quad (3.37)$$

respectively. Let \mathbf{x}_k and \mathbf{y}_l denote the k -th and l -th columns of X and Y , respectively;

$$X = (\mathbf{x}_1, \dots, \mathbf{x}_n), \quad Y = (\mathbf{y}_1, \dots, \mathbf{y}_p). \quad (3.38)$$

Then, Eq. (3.36) means that \mathbf{y}_j , $j = 1, \dots, p$, are eigenvectors of A .

We are in a position to show that $X = YY^T$. Let W denote the eigenspace of X associated with the eigenvalue 1. It follows from $X^2 = X$ and Eq. (3.37) that

$$X(\mathbf{x}_1, \dots, \mathbf{x}_n) = (\mathbf{x}_1, \dots, \mathbf{x}_n), \quad X(\mathbf{y}_1, \dots, \mathbf{y}_p) = (\mathbf{y}_1, \dots, \mathbf{y}_p), \quad (3.39)$$

which means that each column of X and Y is an eigenvector of X associated with the eigenvalue 1. Since $\dim(W) = \text{rank}(X) = p$, we can choose $\mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p}$ so that both $\{\mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p}\}$ and $\{\mathbf{y}_1, \dots, \mathbf{y}_p\}$ may be bases of W . Let Z_0 denote the $n \times p$ full-rank matrix $(\mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_p})$. We decompose Z_0 into

$$Z_0 = QR, \quad Q \in \text{St}(p, n), \quad R \in S_{\text{upp}}^+(p). \quad (3.40)$$

Since R is invertible, the linear span of the column vectors of Q is also W . Thus, there exists a $p \times p$ invertible matrix G such that $Y = QG$. The G proves to be a $p \times p$ orthogonal matrix from Lemma 2.3. Since $QQ^T = X$ from Prop. 3.1, we verify that

$$YY^T = (QG)(QG)^T = QGG^T Q^T = QQ^T = X. \quad (3.41)$$

Conversely, assume that X is expressed as

$$X = YY^T, \quad Y \in \text{St}(p, n), \quad (3.42)$$

where the j -th column \mathbf{y}_j of Y is a normalized eigenvector associated with the eigenvalue λ_{i_j} . Let Λ denote $\text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p})$. Since $AY = Y\Lambda$, we have

$$\text{grad } F(X) = \text{sym}(AX) - XAX = \text{sym}(Y\Lambda Y^T) - YY^T Y\Lambda Y^T = 0. \quad (3.43)$$

This ends the proof. \square

Now that we have characterized critical points for $F(X) = \text{tr}(AX)/2$, we wish to evaluate the value of F at the critical points in order to determine which critical point assigns a minimum value.

Proposition 3.4 *Let X_* be a global optimal solution to Problem 3.1. Then, X_* is expressed as*

$$X_* = Y_* Y_*^T, \quad Y_* \in \text{St}(p, n), \quad (3.44)$$

where the columns of Y_* are orthonormal eigenvectors of A associated with the smallest p eigenvalues $\lambda_1, \dots, \lambda_p$ among all the eigenvalues with $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_p \leq \lambda_{p+1} \leq \dots \leq \lambda_n$.

Proof Let $X_c \in \text{Grass}(p, n)$ be a critical point of F . From Prop. 3.3, X_c is expressed as

$$X_c = YY^T, \quad Y \in \text{St}(p, n), \quad (3.45)$$

where Y is composed of orthonormal eigenvectors of A , which are associated with p eigenvalues $\lambda_{i_1}, \dots, \lambda_{i_p}$ among n eigenvalues of A ;

$$AY = Y \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p}). \quad (3.46)$$

Then, the F is evaluated at X_c as

$$F(X_c) = \frac{1}{2} \text{tr}(AX_c) = \frac{1}{2} \text{tr}(Y \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p}) Y^T) = \frac{1}{2} \sum_{j=1}^p \lambda_{i_j}. \quad (3.47)$$

In order that $F(X_*)$ be the global minimum, the eigenvalues $\{\lambda_{i_j}\}$ in the right-hand side of (3.47) should be the smallest p eigenvalues. This completes the proof. \square

We here note that $Y_* \in \text{St}(p, n)$ in Prop. 3.4 is not uniquely determined. In fact, the permutations of the column vectors of Y_* and the choice of another basis of the eigenspace associated with degenerate eigenvalues $\lambda_i = \lambda_j$ of A for $1 \leq i < j \leq p$, if such a case occurs, leave the X_* invariant. In contrast with this, $X_* \in \text{Grass}(p, n)$ in Prop. 3.4 is uniquely determined, if $\lambda_p \neq \lambda_{p+1}$. However, if $\lambda_p = \lambda_{p+1}$, then X_* is not uniquely determined. We will discuss later what will happen if $\lambda_p = \lambda_{p+1}$.

We proceed to evaluate the Hessian of $F(X) = \text{tr}(AX)/2$ at the critical points.

Proposition 3.5 *Assume that $\lambda_p \neq \lambda_{p+1}$ in the increasing sequence of eigenvalues $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Let X_* be a unique global optimal solution to Problem 3.1. Then, $\text{Hess } F(X_*)$ is positive-definite.*

Proof We shall show that

$$\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} \geq 0, \quad \xi \in T_{X_*} \text{Grass}(p, n), \quad (3.48)$$

and that

$$\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} = 0 \quad (3.49)$$

if and only if ξ is the zero vector of $T_{X_*} \text{Grass}(p, n)$.

From Eq. (3.24), the $\text{Hess } F(X_*)[\xi]$ with $\xi \in T_{X_*} \text{Grass}(p, n)$ is put in the form

$$\text{Hess } F(X_*)[\xi] = \text{sym}(\xi(A - AX_* - X_*A)). \quad (3.50)$$

Since X_* is a critical point of F , from Prop. 3.3 and the course of the proof, there exist $Y \in \text{St}(p, n)$ and $Y_\perp \in \text{St}(n-p, n)$ such that

$$X_* = YY^T, \quad Y^T Y_\perp = 0, \quad (3.51)$$

$$AY = Y \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p}), \quad AY_\perp = Y_\perp \text{diag}(\lambda_{i_{p+1}}, \dots, \lambda_{i_n}). \quad (3.52)$$

We recall here that any tangent vector $\xi \in T_{X_*} \text{Grass}(p, n)$ is put in the form

$$\xi = \text{sym}(YDY_\perp^T), \quad D = (d_{ij}) \in \mathbb{R}^{p \times (n-p)}, \quad (3.53)$$

as is seen in Prop. 2.1, and further that any critical point X_c satisfies $AX_c = X_c A$, as is seen in Eq. (3.30). Using these facts, we evaluate the left-hand side of Eq. (3.48) to obtain

$$\begin{aligned} \langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} &= \frac{1}{4} \text{tr} \left(AY_\perp D^T D Y_\perp^T - AY D D^T Y^T \right) \\ &= \frac{1}{4} \left(\text{tr} \left(\text{diag}(\lambda_{i_{p+1}}, \dots, \lambda_{i_n}) D^T D \right) - \text{tr} \left(\text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p}) D D^T \right) \right) \\ &= \frac{1}{4} \sum_{k=1}^p \sum_{l=1}^{n-p} (\lambda_{i_{p+l}} - \lambda_{i_k}) d_{kl}^2. \end{aligned} \quad (3.54)$$

Since X_* is the global optimal solution, the decomposition of indices is such that $\{i_1, \dots, i_p\} = \{1, \dots, p\}$ and $\{i_{p+1}, \dots, i_n\} = \{p+1, \dots, n\}$ on account of Prop. 3.4 and the ordering of λ_i . Hence, we have for X_*

$$\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} = \frac{1}{4} \sum_{k=1}^p \sum_{l=1}^{n-p} (\lambda_{p+l} - \lambda_k) d_{kl}^2, \quad (3.55)$$

irrespective of the assumption $\lambda_p \neq \lambda_{p+1}$. Taking into account of the ascending order of λ_k , we obtain

$$\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} \geq \frac{1}{4} \sum_{k=1}^p \sum_{l=1}^{n-p} (\lambda_{p+1} - \lambda_p) d_{kl}^2 \geq 0. \quad (3.56)$$

Since $\lambda_{p+1} - \lambda_p > 0$ on account of $\lambda_p \neq \lambda_{p+1}$, Eq. (3.49) holds if and only if $D = (d_{kl}) = 0$, that is, $\xi = 0$ from (3.53). This completes the proof. \square

Prop. 3.5 and Thm. 3.2 are put together to yield the following proposition.

Proposition 3.6 *Let $X_* \in \text{Grass}(p, n)$ be the global optimal solution referred to in Prop. 3.5. Then there exists a neighborhood U of X_* in $\text{Grass}(p, n)$ such that for all $X_0 \in U$, Algorithm 3.6 generates an infinite sequence $\{X_k\}$ converging quadratically to X_* .*

3.6 Degenerate optimal solutions

In what follows, we look into the case where $\text{Hess } F(X_*)$ is degenerate at a global optimal solution X_* . As is easily verified from (3.55), if the p -th eigenvalue λ_p is of multiplicity r and $\lambda_p = \lambda_{p+1} = \cdots = \lambda_{p+r-1}$, the d_{kl} with $k = p$ and $l = 1, \dots, r-1$ make no contribution to the value of $\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*}$, so that the degenerate subspace of $\text{Hess } F(X_*)$ is considered to be an $(r-1)$ -dimensional subspace of $T_{X_*} \text{Grass}(p, n)$ at least. We show that if A has a degenerate eigenvalue, there exist a continuum of critical points on which the Hessian of F is degenerate, under a more general degeneracy condition than the stated above.

Proposition 3.7 *Let s be a nonnegative integer not larger than p . Assume that a symmetric matrix A has an eigenvalue μ of multiplicity r such that $p-s \leq r \leq n-s$. Let μ_1, \dots, μ_s denote any eigenvalues (not especially ordered) of A which are not equal to any other $n-s$ eigenvalues. Further, denote by $Y \in \text{St}(p, n)$ the matrix formed by the orthonormal eigenvectors and satisfying $AY = Y \text{diag}(\mu_1, \dots, \mu_s, \underbrace{\mu, \dots, \mu}_{p-s})$. Then, the corresponding critical points $X_c = YY^T \in$*

$\text{Grass}(p, n)$ of F in Problem 3.1 form a submanifold diffeomorphic to the Grassmann manifold $\text{Grass}(p-s, r)$. We identify the submanifold with $\text{Grass}(p-s, r)$ and denote the inclusion map by $\iota : \text{Grass}(p-s, r) \hookrightarrow \text{Grass}(p, n)$. Further, for each point P of $\text{Grass}(p-s, r)$, the Hessian $\text{Hess } F(X_c)$ at $X_c = \iota(P)$ is degenerate for the tangent space $T_P \text{Grass}(p-s, r)$ viewed as a subspace of $T_{X_c} \text{Grass}(p, n)$.

Proof Let μ_1, \dots, μ_n be eigenvalues of A , where $\mu_{s+1} = \cdots = \mu_p = \cdots = \mu_{s+r} = \mu$. We put Y in the form $Y = (v_1, \dots, v_s, v'_{s+1}, \dots, v'_p)$, where $v_1, \dots, v_s, v'_{s+1}, \dots, v'_p$ are orthonormal eigenvectors of A associated with the eigenvalues $\mu_1, \dots, \mu_s, \mu_{s+1}, \dots, \mu_p$, respectively, where $\mu_{s+1} = \cdots = \mu_p = \mu$. Let $\{v_{s+1}, \dots, v_{s+r}\}$ denote a fixed orthonormal basis of the eigenspace associated with μ . Then, the v'_{s+j} , $j = 1, \dots, p-s$, can be expressed as $v'_{s+j} = \sum_{k=1}^r q_{kj} v_{s+k}$, where the matrix $Q = (q_{ij})$ is an $r \times (p-s)$ orthonormal matrix. Let v_{s+r+1}, \dots, v_n denote orthonormal eigenvectors of A associated with eigenvalues $\mu_{s+r+1}, \dots, \mu_n$, respectively, and V denote the matrix formed by the orthonormal eigenvectors $\{v_1, \dots, v_n\}$; $V = (v_1, \dots, v_n) \in O(n)$. Then, Y can be put in the form

$$Y = (v_1, \dots, v_s, v'_{s+1}, \dots, v'_p) = (v_1, \dots, v_n) \begin{pmatrix} I_s & 0 \\ 0 & Q \\ 0 & 0 \end{pmatrix} = V \begin{pmatrix} I_s & 0 \\ 0 & Q \\ 0 & 0 \end{pmatrix}, \quad (3.57)$$

and thereby the corresponding critical point X_c is expressed as

$$X_c = YY^T = V \begin{pmatrix} I_s & 0 & 0 \\ 0 & QQ^T & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T. \quad (3.58)$$

Since $Q \in \text{St}(p-s, r)$, the quantity $P = QQ^T$ belongs to $\text{Grass}(p-s, r)$. Hence, Eq. (3.58) implies that the critical points $\iota(P) = X_c$ form a submanifold diffeomorphic to $\text{Grass}(p-s, r)$.

Let $Q_\perp \in \text{St}(r-p+s, r)$ denote a matrix whose columns are orthonormal vectors to the column vectors of Q in \mathbb{R}^r , and $Y_\perp \in \text{St}(n-p, n)$ denote the matrix

$$Y_\perp = V \begin{pmatrix} 0 & 0 \\ Q_\perp & 0 \\ 0 & I_{n-s-r} \end{pmatrix}. \quad (3.59)$$

Then, it follows from Prop. 2.1 that any tangent vector $\xi \in T_{X_c} \text{Grass}(p, n)$ is expressed as $\xi = \text{sym}(YDY_\perp^T)$, where $D = (d_{ij}) \in \mathbb{R}^{p \times (n-p)}$. From Eq. (3.54), the quadratic form $\langle \text{Hess } F(X_c)[\xi], \xi \rangle_{X_c}$ is evaluated as

$$\begin{aligned} & \langle \text{Hess } F(X_c)[\xi], \xi \rangle_{X_c} \\ &= \frac{1}{4} \sum_{k=1}^p \sum_{l=1}^{n-p} (\mu_{p+l} - \mu_k) d_{kl}^2 \\ &= \frac{1}{4} \left(\sum_{k=1}^s \sum_{l=1}^{n-p} (\mu_{p+l} - \mu_k) d_{kl}^2 + \sum_{k=s+1}^p \sum_{l=1}^{s+r-p} (\mu_{p+l} - \mu_k) d_{kl}^2 + \sum_{k=s+1}^p \sum_{l=s+r-p+1}^{n-p} (\mu_{p+l} - \mu_k) d_{kl}^2 \right). \end{aligned} \quad (3.60)$$

According to the right-hand side of (3.60), we partition D into

$$D = \begin{pmatrix} D_1 & D_2 \\ D_3 & D_4 \end{pmatrix}, \quad (3.61)$$

where $D_1 \in \mathbb{R}^{s \times (s+r-p)}$, $D_2 \in \mathbb{R}^{s \times (n-s-r)}$, $D_3 \in \mathbb{R}^{(p-s) \times (s+r-p)}$, and $D_4 \in \mathbb{R}^{(p-s) \times (n-s-r)}$. It then follows from the conditions $\mu_{s+1} = \dots = \mu_p = \dots = \mu_{s+r} = \mu$ and $\mu_1, \dots, \mu_s, \mu_{s+r+1}, \dots, \mu_n \neq \mu$ that $\langle \text{Hess } F(X_c)[\xi], \xi \rangle_{X_c} = 0$ only for $\xi = \text{sym}(YDY_\perp^T)$ with $D_1 = 0$, $D_2 = 0$, and $D_4 = 0$. This implies in general that the degeneracy subspace is isomorphic to $\mathbb{R}^{(p-s) \times (s+r-p)}$ corresponding to D_3 and spanned by

$$\begin{aligned} \xi &= \text{sym}(YDY_\perp^T) \\ &= \text{sym} \left(V \begin{pmatrix} I_s & 0 \\ 0 & Q \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ D_3 & 0 \end{pmatrix} \begin{pmatrix} 0 & Q_\perp^T & 0 \\ 0 & 0 & I_{n-s-r} \end{pmatrix} V^T \right) \\ &= V \begin{pmatrix} 0 & 0 & 0 \\ 0 & \text{sym}(QD_3Q_\perp^T) & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T, \end{aligned} \quad (3.62)$$

with D_3 varying in $\mathbb{R}^{(p-s) \times (s+r-p)}$. Eq. (3.62) is viewed as the tangent map $T_{QQ^T} \text{Grass}(p-s, r) \rightarrow T_{X_c} \text{Grass}(p, n)$ to the inclusion map $\text{Grass}(p-s, r) \rightarrow \text{Grass}(p, n)$ defined by (3.58). Here, we note that $\left\{ \text{sym}(QD_3Q_\perp^T) \mid D_3 \in \mathbb{R}^{(p-s) \times (s+r-p)} \right\} \simeq T_{QQ^T} \text{Grass}(p-s, r)$, as is seen from Eq. (2.3). This completes the proof. \square

If $X_c = X_*$ is a global optimal solution in Prop. 3.7, Eq. (3.60) implies that $\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} = 0$ if and only if ξ is of the form (3.62). Thus, we have the following corollary.

Corollary 3.2 *Assume that A has the eigenvalues $\lambda_1 \leq \cdots \leq \lambda_s < \lambda_{s+1} = \cdots = \lambda_p = \cdots = \lambda_{s+r} < \lambda_{s+r+1} \leq \cdots \leq \lambda_n$, where $0 \leq s \leq p$ and $p - s \leq r \leq n - s$. Then, the set of global optimal solutions X_* in Problem 3.1 forms a submanifold diffeomorphic to $\text{Grass}(p - s, r)$. Further, the Hessian $\text{Hess } F(X_*)$ at each X_* is degenerate and $\langle \text{Hess } F(X_*)[\xi], \xi \rangle_{X_*} = 0$ if and only if ξ is tangent to the submanifold stated above.*

3.7 Newton's equation at a degenerate critical point

We make a remark on the implication of the degenerate Hessian for Newton's equation (3.23) at a critical point X_c of F in Problem 3.1. We show that the degeneracy of the Hessian is closely related to the non-uniqueness of solutions to Newton's equation (3.23). For this purpose, we look into the eigenvalues of the matrix B included in (3.23).

Proposition 3.8 *Let μ_1, \dots, μ_n denote the eigenvalues of A , which are not especially ordered. Assume that X_c is a critical point of F in Problem 3.1, which is put in the form*

$$X_c = YY^T, \quad Y = (u_1, \dots, u_p) \in \text{St}(p, n), \quad (3.63)$$

where u_1, \dots, u_n are orthonormal eigenvectors associated with μ_1, \dots, μ_n , respectively. Then the eigenvalues of the matrix $B = A - AX_c - X_cA$ are $-\mu_1, \dots, -\mu_p, \mu_{p+1}, \dots, \mu_n$.

Proof Since X_c is a critical point, Eq. (3.30) holds, so that the matrix B is brought into $B = A - 2AX_c$. We operate the basis vectors $\{u_i\}_{i=1, \dots, n}$ with B . For $1 \leq i \leq p$, we have

$$Bu_i = Au_i - 2AYY^T u_i = Au_i - 2Au_i = -\mu_i u_i, \quad (3.64)$$

and for $p + 1 \leq j \leq n$,

$$Bu_j = Au_j - 2AYY^T u_j = Au_j = \mu_j u_j. \quad (3.65)$$

These equations show that the eigenvalues of B are $-\mu_1, \dots, -\mu_p, \mu_{p+1}, \dots, \mu_n$. \square

If X_* is a global optimal solution, then X_* is put in the form

$$X_* = YY^T, \quad Y = (v_1, \dots, v_p), \quad (3.66)$$

where v_1, \dots, v_p are orthonormal eigenvectors associated with the smallest p eigenvalues $\lambda_1, \dots, \lambda_p$, respectively. Applying Prop. 3.8 in this case, we obtain the following corollary.

Corollary 3.3 *For a global optimal solution X_* to Problem 3.1, the eigenvalues of the matrix $B = A - AX_* - X_*A$ are $-\lambda_1, \dots, -\lambda_p, \lambda_{p+1}, \dots, \lambda_n$.*

With this in mind, we look at the Lyapunov equation

$$\eta B + B\eta = C, \quad B, C \in \mathbb{R}^{n \times n} \quad (3.67)$$

for $\eta \in \mathbb{R}^{n \times n}$, where B is a symmetric matrix. As is well known, Eq. (3.67) has a unique solution if and only if B and $-B$ have no common eigenvalues [9]. We apply this claim to the Lyapunov equation at a global optimal solution X_* to Problem 3.1 with doubly degenerate eigenvalues $\lambda_p = \lambda_{p+1}$ of A . Since the matrix $B = A - AX_* - X_*A$ has eigenvalues $-\lambda_p$ and λ_{p+1} , B and $-B$ have the common eigenvalue $-\lambda_p = -\lambda_{p+1}$. This implies that the Lyapunov equation (3.67) has no unique solution in $\mathbb{R}^{n \times n}$. The non-uniqueness of solutions to the Lyapunov equation in the case of $\lambda_p = \lambda_{p+1}$ is consistent with the degeneracy of the Hessian at X_* . In fact, from Cor. 3.2, $\text{Hess } F(X_*)$ is degenerate, so that Newton's equation, $\text{Hess } F(X_*)[\eta] = -\text{grad } F(X_*)$, which is put in the form

$$\eta B + B\eta = 0, \quad \eta \in T_{X_*} \text{Grass}(p, n), \quad (3.68)$$

has indeed a non-trivial solution $\eta \neq 0$. Here, we note that we have distinguished Eq. (3.68) from (3.67); η is considered to be in $\mathbb{R}^{n \times n}$ for the Lyapunov equation (3.67) and in $T_{X_*} \text{Grass}(p, n)$ for Newton's equation (3.68), where $T_{X_*} \text{Grass}(p, n)$ is a linear subspace of $T_{X_*} \mathbb{R}^{n \times n} \simeq \mathbb{R}^{n \times n}$.

3.8 Computation for degenerate optimal solutions

In practice, however, no matter whether $\lambda_p = \lambda_{p+1}$ or not, Algorithm 3.6 works well. We observe what happens in a sufficiently small neighborhood U of a global optimal solution X_* . If $\lambda_p \neq \lambda_{p+1}$, then the Hessian $\text{Hess } F(X_*)$ at X_* is non-degenerate, so that $\text{Hess } F(X)$ is also non-degenerate for any $X \in U$ sufficiently close to X_* . Then, Newton's equation (3.23) with $X \in U$ has a unique solution. Once the computer at work decides that the current iterate X_k is equal to X_* , that is, the error is less than the machine epsilon, then C , twice the gradient of F at X_k , is evaluated as 0 and the next Newton's direction η_k is calculated as $\eta_k = 0$, so that $X_{k+1} = X_k$. This implies that Algorithm 3.6 no longer improves the current iterate and the global optimal solution is obtained within the accuracy of the machine epsilon. If $\lambda_p = \lambda_{p+1}$, then neither the Lyapunov equation (3.67) with $C = 0$ nor Newton's equation (3.68) can be uniquely solved at any global optimal solution X_* , and the algorithm for solving Lyapunov equations stops automatically. In such a case, however, a global solution X_* is already obtained as the current iterate X_k within the accuracy of the machine epsilon. We recall here that global optimal solutions in this case are not unique but form a submanifold. Thus, we can conclude that in both cases, we can obtain a global optimal solution X_* through Algorithm 3.6, if the initial point X_0 is sufficiently close to X_* .

To show that Newton's method yields indeed a set of global optimal solutions when $\lambda_p = \lambda_{p+1}$, we take an example with $n = 3$, $p = 1$, $A = \text{diag}(1, 1, 2)$. In this case, one has $s = 0$ and $r = 2$ since $\lambda_1 = \lambda_2 < \lambda_3$. From Eq. (3.58), a global optimal solution X_* is put in the form

$$X_* = \begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}, \quad (3.69)$$

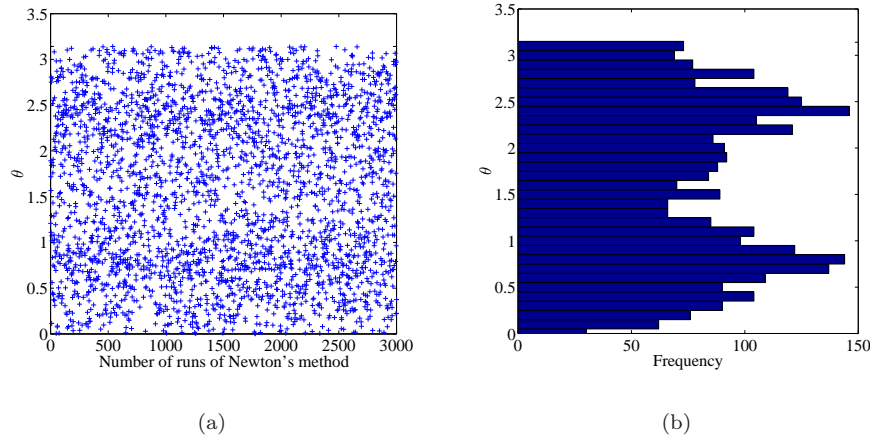


Fig. 3.1 Values of θ corresponding to P in Eq. (3.69) with 3×10^3 different initial points.

where the 2×2 matrix P belongs to $\text{Grass}(1, 2) = \mathbb{RP}^1$. For each $P \in \mathbb{RP}^1$, there exists a unique $q = (\cos \theta, \sin \theta)^T \in \text{St}(1, 2) = S^1$ with $0 \leq \theta < \pi$ such that $P = qq^T$. Put another way, there is a one-to-one correspondence between $X_* \in \text{Grass}(1, 3)$ and $\theta \in [0, \pi)$. Such a θ can be calculated as follows: From Prop. 3.1, $q_0 = (\cos \phi, \sin \phi)^T$, $0 \leq \phi < 2\pi$ satisfying $P = q_0 q_0^T$ can be obtained by the QR decomposition of a non-zero column vector of P . Then, $\theta = \phi$ if $0 \leq \phi < \pi$ and $\theta = \phi - \pi$ if $\pi \leq \phi < 2\pi$. Performing Newton's method repeatedly for various randomly chosen initial points on $\text{Grass}(1, 3)$, we obtain solutions of the form (3.69) and then the corresponding values of θ . Plotting the values of θ for each run of Newton's method results in the left figure (a) of Fig. 3.1. There are 3000 dots in Fig. 3.1(a), which are counted in each subinterval of θ to yield the right figure (b) as a histogram of global optimal solutions obtained by Newton's method. The figure shows that sequences generated by Newton's method with different initial values converge to different limit points which fill up the interval $[0, \pi)$ for θ . These limit points indeed form a submanifold $\mathbb{RP}^1 = \text{Grass}(1, 2)$ of $\text{Grass}(1, 3)$.

4 A hybrid method for the Rayleigh quotient on $\text{Grass}(p, n)$

4.1 Numerical observations and analysis

We perform Algorithm 3.6 for the function $F(X) = \text{tr}(AX)/2$ on $\text{Grass}(p, n)$ with $n = 3$, $p = 1$, $A = \text{diag}(1, 2, 3)$. The variation of the value $F(X)$ along a generated sequence $\{X_k\}$ is shown in Fig. 4.1, where k ranges from 0 to 50. The vertical axis of Fig. 4.1 carries the values of F along $\{X_k\}$.

Since Eq. (3.47) implies that the value of F at a critical point is half of the sum of p eigenvalues among n eigenvalues of A , possible critical values $F(X_c)$ in the present case are 0.5, 1.0, 1.5. Fig. 4.1 shows that the sequence $\{X_k\}$ generated by Algorithm 3.6 approaches critical points, but if a temporary target critical point

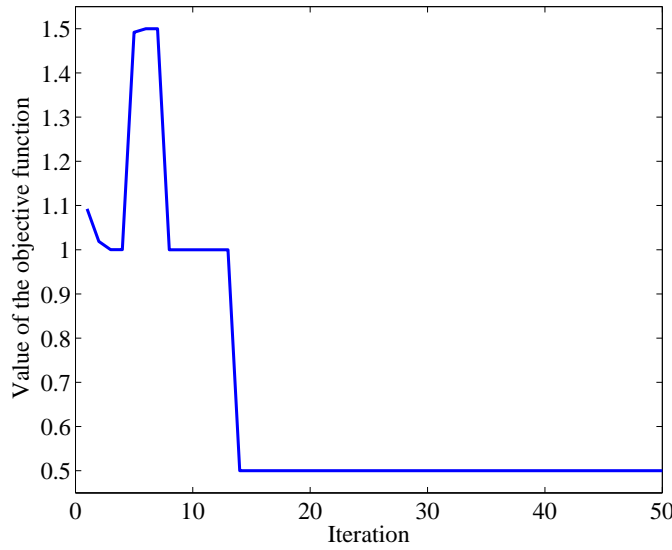


Fig. 4.1 $n = 3, p = 1, A = \text{diag}(1, 2, 3)$.

is not the global minimizer $\text{diag}(1, 0, 0)$, then X_k goes out to other critical points after a while. This happens partly because A is diagonal. Indeed, when a point $X_k \in \text{Grass}(1, 3)$ sufficiently close to a critical point $\text{diag}(0, 1, 0)$ is obtained, the leftmost column \mathbf{x}_1 of X_k is nearly equal to 0. If the computer decides that \mathbf{x}_1 is equal to 0, then Y_k is correctly computed as $Y_k = \text{qf}(\mathbf{x}_2)$ by Algorithm 3.2. However, if the computer decides that $\mathbf{x}_1 \neq 0$ and uses it to compute $Y_k = \text{qf}(\mathbf{x}_1)$, the resultant Y_k is not a correct one. In other words, $Y_k Y_k^T$ is no longer equal to X_k , and the exponential or the QR-based retraction is performed with the wrong Y_k to produce the next iterate X_{k+1} . This means that no matter how small is the deviation of X_k from $\text{Grass}(1, 3)$, the computational error contained in \mathbf{x}_1 is

fatal. For example, for $X_k = \begin{pmatrix} -10^{-19} & -10^{-21} & -10^{-17} \\ -10^{-21} & 1 & 10^{-19} \\ -10^{-17} & 10^{-19} & -10^{-16} \end{pmatrix}$, which approximates to

$\text{diag}(0, 1, 0)$, the computer uses $\mathbf{x}_1 = (-10^{-19}, -10^{-21}, -10^{-17})^T$ to determine $Y_k = \text{qf}(\mathbf{x}_1)$. Since the absolute value of the third component of \mathbf{x}_1 is 10^2 times as large as the second largest one, the evaluated Y_k approximates to $(0, 0, -1)^T$. Then, the corresponding $Y_k Y_k^T$ is not equal to the given X_k but nearly equal to another critical point $\text{diag}(0, 0, 1)$. For this X_k , the matrix Frobenius norm $\|X_k^2 - X_k\|_F$ of $X_k^2 - X_k$ (which is supposed to be 0 on account of the constraint $X_k^2 = X_k$) is about 10^{-16} . Since the error is at the level of 10^{-16} , the computer should decide that $\mathbf{x}_1 = (-10^{-19}, -10^{-21}, -10^{-17})^T$ is equal to 0 with the same level of accuracy. However, the computer uses \mathbf{x}_1 to determine Y_k , if we do not require the order of accuracy for \mathbf{x}_1 in advance. Since the next iterate X_{k+1} is located in the neighborhood of $\text{diag}(0, 0, 1)$, the value of F jumps upward at X_{k+1} . If the first component of \mathbf{x}_1 is the largest in absolute value, then $Y_k = \text{qf}(\mathbf{x}_1)$ approximates

to $(1, 0, 0)^T$, so that the next iterate X_{k+1} is close to $\text{diag}(1, 0, 0)$, and thereby the value of F jumps downward at X_{k+1} . These jumps are observed in Fig. 4.1.

We can avoid this mistake in determining Y_k as follows: Reordering the diagonal elements of A , we consider the case $n = 3$, $p = 1$, $A = \text{diag}(2, 1, 3)$. Then, any sequence $\{X_k\}$ generated by Algorithm 3.6 with initial points close to the critical point $\text{diag}(1, 0, 0)$ corresponding to the middle critical value $F(X_c) = 1$ stays in the neighborhood of the same critical point $\text{diag}(1, 0, 0)$. In this example, no problem occurs for determining Y_k .

As is mentioned already, a way to avoid such a mistake in determining Y_k is to ignore small numbers within the accuracy required for the sequence X_k . In fact, in the same example as in Fig. 4.1, if we ignore numbers less than 10^{-16} , we get a sequence which does not wander from a critical point to another.

The issue comes from the fact that it is difficult to determine the linear dependence/independence among vectors in numerical manner. In effect, if we take a symmetric matrix A randomly, then the wandering of $F(X_k)$ such as in Fig. 4.1 rarely appears. This is because the leftmost p columns of a critical point is often linearly independent. On the contrary, the case in which A is diagonal is rather a rare case in which the leftmost p columns of a critical point is not linearly independent.

We move to another example with $n = 10$, $p = 5$, $A = Q \text{diag}(1, 2, \dots, 10)Q^T$, where Q is a 10×10 orthogonal matrix with randomly chosen elements. Fig. 4.2

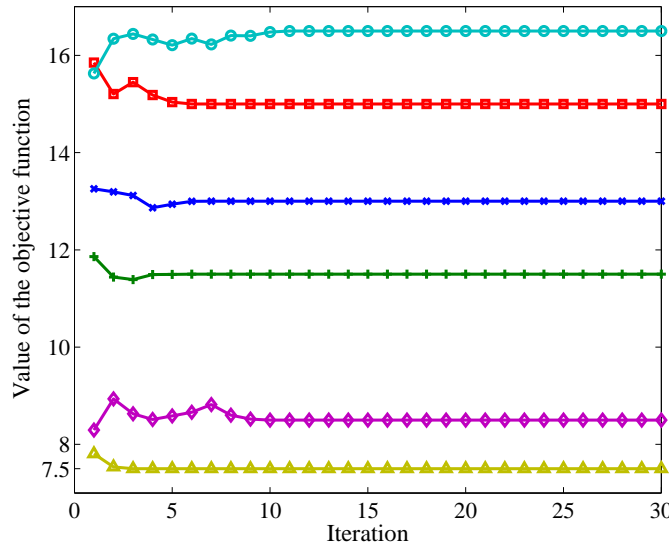


Fig. 4.2 $n = 10, p = 5$, A is a symmetric matrix whose eigenvalues are $1, 2, \dots, 10$.

shows the values of F along sequences $\{X_k\}$ generated by Algorithm 3.6 with different initial points. Once a sequence arrives in the convergence region of a temporary target critical point X_c , the values $F(X_k)$ go up and down to tend

to $F(X_c)$. In contrast with this, if the sequence starts in the neighborhood of a global minimum point X_* , then the values $F(X_k)$ are monotonically decreasing to $F(X_*) = 7.5$.

In order to understand these behaviors of $F(X_k)$, we study the properties of critical points by resorting to the Hessian at critical points.

Proposition 4.1 *Assume that A is an $n \times n$ symmetric matrix with distinct eigenvalues $\lambda_1 < \dots < \lambda_n$. Let X_c be a critical point of the objective function $F(X) = \text{tr}(AX)/2$ on $\text{Grass}(p, n)$. If X_c is neither the global maximum point nor the global minimum point of F , then X_c is a saddle point of F .*

Proof For $\xi = \text{sym}(YDY_\perp^T) \in T_{X_c}\text{Grass}(p, n)$, from Eq. (3.54), we have

$$\langle \text{Hess } F(X_c)[\xi], \xi \rangle_{X_c} = \frac{1}{4} \sum_{k=1}^p \sum_{l=1}^{n-p} (\lambda_{i_{p+l}} - \lambda_{i_k}) d_{kl}^2. \quad (4.1)$$

If X_c is neither a global maximizer nor a global minimizer, each of the coefficients $\lambda_{i_{p+l}} - \lambda_{i_k}$ can take a negative or a positive value. Thus, $\text{Hess } F(X_c)$ is indefinite as a symmetric operator on $T_{X_c}\text{Grass}(p, n)$. This completes the proof. \square

Let X_c be a critical point other than a maximizer and a minimizer. Then, for a solution η to Newton's equation $\text{Hess } F(X_k)[\eta] = -\text{grad } F(X_k)$, one has

$$\langle \eta, \text{Hess } F(X_k)[\eta] \rangle_{X_k} = -\langle \eta, \text{grad } F(X_k) \rangle_{X_k} = -DF(X_k)[\eta]. \quad (4.2)$$

If X_k is sufficiently close to X_c , $\text{Hess } F(X_k)$ is also close to $\text{Hess } F(X_c)$. Therefore, if $\langle \eta, \text{Hess } F(X_k)[\eta] \rangle_{X_k}$ is negative, then $DF(X_k)[\eta]$ is positive, so that the value of F is increasing in Newton's direction η . On the contrary, if $\langle \eta, \text{Hess } F(X_k)[\eta] \rangle_{X_k}$ is positive, then the value of F is decreasing in Newton's direction η . Thus the values $F(X_k)$ may go upward or downward along $\{X_k\}$, depending on where X_k is placed in a neighborhood of X_c , and eventually tend to the target value $F(X_c)$. Fig. 4.2 gives an example of such variations along $\{X_k\}$.

To show these behaviors explicitly, we present an exact example. Let $n = 3$, $p = 1$, $A = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$, where $\lambda_1 < \lambda_2 < \lambda_3$. In this case, the point $X_c = \text{diag}(0, 1, 0)$ in $\text{Grass}(1, 3)$ is a saddle point. Let $X \in \text{Grass}(1, 3)$ be a point sufficiently near to X_c . Then, there exist $\xi \in T_{X_c}\text{Grass}(1, 3)$ and a sufficiently small $t_0 > 0$ such that

$$X = \text{Exp}_{X_c}(t_0 \xi), \quad (4.3)$$

where the exponential map Exp is given by Eq. (2.37). It follows from Prop. 2.1 that $\xi \in T_{X_c}\text{Grass}(1, 3)$ can be expressed as

$$\xi = \begin{pmatrix} 0 & a & 0 \\ a & 0 & b \\ 0 & b & 0 \end{pmatrix}, \quad a, b \in \mathbb{R}. \quad (4.4)$$

We may set $\|\xi\|_{X_c} = \sqrt{2}$, so that $a^2 + b^2 = 1$. Then, putting $a = \cos \phi$, $b = \sin \phi$, where $0 \leq \phi < 2\pi$, we can assign X by two parameters t_0 and ϕ . Let $\eta \in T_X\text{Grass}(1, 3)$ be the solution to Newton's equation

$$\eta(A - AX - XA) + (A - AX - XA)\eta = 2(-\text{sym}(AX) + XAX). \quad (4.5)$$

Since Eq. (4.5) is a linear equation, we can write out the components of η and evaluate the quantity $DF(X)[\eta] = \text{tr}(A\eta)/2$ specifically. From a straightforward calculation, it turns out that for a sufficiently small t_0 , the sign of $DF(X)[\eta]$ is independent of t_0 . Put in detail, if $0 \leq \phi < \alpha$, $\pi - \alpha < \phi < \pi + \alpha$, or $2\pi - \alpha < \phi < 2\pi$, then $DF(X)[\eta] < 0$, and further if $\alpha < \phi < \pi - \alpha$ or $\pi + \alpha < \phi < 2\pi - \alpha$, then $DF(X)[\eta] > 0$, where $\alpha = \arccos((\Lambda_3 + \Lambda_1)/(\Lambda_3 - \Lambda_1))/2$, $\Lambda_1 = \lambda_1 - \lambda_2$, $\Lambda_3 = \lambda_3 - \lambda_2$. This implies that in any neighborhood of the saddle point X_c , there exist two types of open subsets at one of which the function F is decreasing in Newton's direction and at the other of which the F is increasing in Newton's direction.

4.2 A hybrid method

In contrast with Newton's method, the steepest descent method is a descent algorithm, so that only the global minimum point is stable and the other critical points unstable. We should adopt the steepest descent method to prevent the sequence $\{X_k\}$ from being trapped by a non-minimum critical point. However, the speed of the convergence to an optimal solution in the steepest descent method is less than that in Newton's method, if the sequence is in a vicinity of the optimal solution. To speed up the convergence, we propose a hybrid method for Problem 3.1. The strategy is as follows: We first perform the steepest descent method until $\text{grad } F(X_k)$ is sufficiently small. Then, we switch the method to Newton's one. The algorithm is described as follows:

Algorithm 4.1 Hybrid method for Problem 3.1

- 1: Choose an initial point $X_0 \in \text{Grass}(p, n)$ and a parameter $\varepsilon > 0$. Set $k := 0$.
 - 2: **while** $\sqrt{\text{tr}((\text{sym}(AX_k) - X_k AX_k)^2)} > \varepsilon$ **do**
 - 3: Perform Step 3-5 in Algorithm 3.5.
 - 4: $k := k + 1$.
 - 5: **end while**
 - 6: Set $X_0 := X_k$ and $k := 0$.
 - 7: Perform Algorithm 3.6.
-

We have to make a remark on global convergence of the sequence $\{X_k\}$. Though critical points other than the global optimal solution are unstable with respect to the gradient flow, each critical point has its stable manifold (see [1] for the definition, for example). By definition, this implies that there is a set of initial points with which sequences generated by the steepest descent method start to converge to a saddle point. However, such a set is of dimension less than that of the whole space. For this reason, we obtain the following proposition.

Proposition 4.2 *For almost all initial point $X_0 \in \text{Grass}(p, n)$, Algorithm 4.1 with sufficiently small ε generates a sequence globally quadratically converging to a global optimal solution to Problem 3.1.*

We will discuss stable manifolds for a critical point of $F(X)$ in detail later in Subsection 4.4 and further give a sufficient condition for the sequence generated by

the steepest descent method algorithm to converge to a saddle point in Subsection 4.5.

We give a numerical example of the hybrid method. Algorithm 4.1 with $\varepsilon = 0.5$ is applied to the function $F(X) = \text{tr}(AX)/2$ with $A = P \text{diag}(1, \dots, n)P^T$ of several different sizes, where P is a randomly chosen orthogonal matrix. The numerical results after switching to Newton's method are shown in the following table, where k denotes the iteration number of Newton's method which is applied on Step 7 in Algorithm 4.1, and where the error of X_k from the optimal solution X_* is measured by $\|\text{grad } F(X_k)\|_{X_k}$.

Table 4.1 Error of X_k from the optimal solution for several A with different sizes.

$(n, p) \backslash k$	0	1	2	3	4
(50, 10)	0.4829	$10^{-0.9727}$	$10^{-3.0789}$	$10^{-9.4099}$	$10^{-22.6250}$
(50, 30)	0.4505	$10^{-1.1066}$	$10^{-3.4908}$	$10^{-10.6477}$	$10^{-23.6657}$
(100, 10)	0.4996	$10^{-0.9075}$	$10^{-2.8419}$	$10^{-8.6789}$	$10^{-21.7203}$
(100, 30)	0.4712	$10^{-1.0242}$	$10^{-3.2385}$	$10^{-9.8892}$	$10^{-23.4476}$
(100, 50)	0.4949	$10^{-0.8088}$	$10^{-2.5563}$	$10^{-7.8362}$	$10^{-21.2276}$
(100, 70)	0.4975	$10^{-0.9127}$	$10^{-2.8970}$	$10^{-8.8639}$	$10^{-21.9213}$
(100, 90)	0.4967	$10^{-0.9594}$	$10^{-3.0486}$	$10^{-9.3218}$	$10^{-21.6299}$
(300, 150)	0.4983	$10^{-0.9533}$	$10^{-3.0298}$	$10^{-9.2654}$	$10^{-21.3901}$

From Table 4.1, we observe that once X_k arrives in the appropriate region determined by ε , Newton's method switched from the steepest descent method generates sequences quickly converging.

If we take a matrix A of the form $A = P \text{diag}(1, \dots, n)P^T$ with P a randomly chosen orthogonal matrix and set $(n, p) = (300, 150)$, and if we adopt the hybrid method, we obtain numerical results which are arranged in Fig. 4.3. This figure shows that the switching of the steepest descent method to Newton's one accelerates drastically the speed of convergence: The upper sequence of dots shows that the convergence of the sequence generated by the steepest descent method is very slow. If the sequence generating method is switched to Newton's method after the iterate X_k gets into the ε -neighborhood of the target, the generated sequence reaches, in a few steps, the final point within the computer accuracy (see two dots in the middle of the figure), and no update is seen thereafter (see the bottom sequence of dots).

4.3 Measurement of accuracy

Our present method has another merit as an eigenvalue algorithm. Since Newton's method with an appropriate initial point generates a sequence converging quadratically to an optimal solution on account of Prop. 3.6, we can use our method to obtain a more accurate solution than that obtained by the existing eigenvalue algorithms. Suppose we have an approximate solution by using an existing algorithm. We take it as an initial point of our Newton's method. Performing our method, we obtain a more accurate solution. The accuracy can

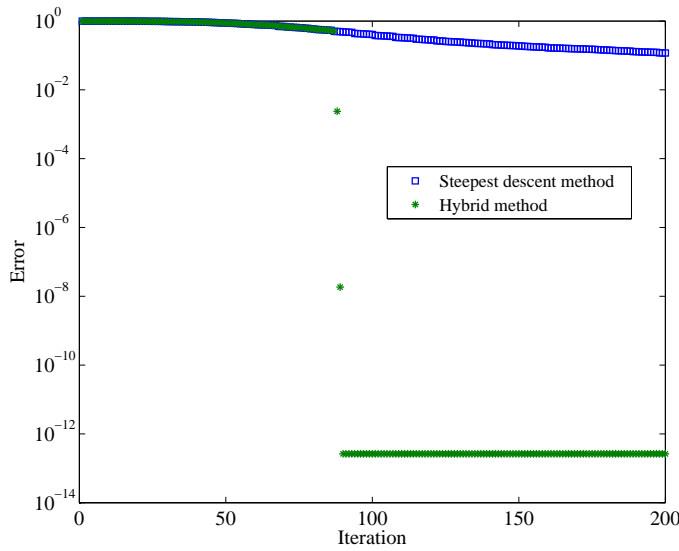


Fig. 4.3 $n = 300, p = 150$.

be verified by observing that the sequence generated by Newton's method reduces the value of the objective function. For example, we consider Problem 3.1 with $n = 700$, $p = 500$, and $A = P \text{diag}(1, 2, \dots, 700)P^T$, where P is a randomly chosen orthogonal matrix. The optimal solution X_{opt} to this problem is $X_{\text{opt}} = Y_{\text{opt}}Y_{\text{opt}}^T$, where $Y_{\text{opt}} = P \begin{pmatrix} I_p \\ 0 \end{pmatrix}$. We perform MATLAB's eigs function to obtain $Y_{\text{ML}} \in \text{St}(p, n)$, whose i -th column vector from the left is an (approximate) eigenvector associated with the i -th smallest eigenvalue of A . Set $X_{\text{ML}} := Y_{\text{ML}}Y_{\text{ML}}^T$, which is of course a good approximate solution of Problem 3.1. We can measure the difference between X_{opt} and X_{ML} by taking the value of the objective function F at each point. The result is

$$\frac{1}{2} \text{tr}(AX_{\text{ML}}) - \frac{1}{2} \text{tr}(AX_{\text{opt}}) = 3.6380 \times 10^{-11}, \quad (4.6)$$

which shows that the solution X_{ML} admits of improvement. We here set $X_0 := X_{\text{ML}}$ and perform Algorithm 3.6 to obtain X_{New} . Then, we have the result

$$\frac{1}{2} \text{tr}(AX_{\text{ML}}) - \frac{1}{2} \text{tr}(AX_{\text{New}}) = 3.6380 \times 10^{-11} > 0, \quad (4.7)$$

which shows that the solution X_{New} is an improved solution in comparison with X_{ML} . Moreover, we have

$$\frac{1}{2} \text{tr}(AX_{\text{New}}) - \frac{1}{2} \text{tr}(AX_{\text{opt}}) = 0 \quad (4.8)$$

on the computer, which means that X_{New} is as accurate solution as the computer can obtain.

If we do not know an exact optimal solution, we can use the value of the objective function at the final point in order to compare the accuracy among approximate solutions obtained by various existing methods.

4.4 Convergence to a saddle point

We return to a question of global convergence, as we mentioned before. The sequence $\{X_k\}$ generated by the steepest descent method may not necessarily converge to the global optimal solution but to a saddle point X_c if the initial point X_0 is in the stable manifold of X_c with respect to the gradient flow. However, the stable manifold is thin. In fact, for a saddle point X_c , the tangent space to the stable manifold of X_c is explicitly described as a dimension-deficient vector space, as is shown below.

Proposition 4.3 *Let $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ be the increasing sequence of eigenvalues of A . Let $X_c = YY^T \in \text{Grass}(p, n)$ be a critical point of $F(X) = \text{tr}(AX)/2$ with $AY = Y \text{diag}(\lambda_{i_1}, \dots, \lambda_{i_p})$, where $i_1, \dots, i_p \in \{1, \dots, n\}$ are distinct integers with $i_1 < \dots < i_p$. Denote by $\{i_{p+1}, \dots, i_n\}$ the complement of the subset $\{i_1, \dots, i_p\}$ in $\{1, \dots, n\}$ with $i_{p+1} < \dots < i_n$. Further, let I_s and I_u be the sets of pairs of indices defined by $I_s := \{(k, l) \in \mathbb{N}^2 \mid i_k < i_{p+l}, \lambda_{i_k} \neq \lambda_{i_{p+l}}\}$ and $I_u := \{(k, l) \in \mathbb{N}^2 \mid i_k > i_{p+l}, \lambda_{i_k} \neq \lambda_{i_{p+l}}\}$, respectively. Consider the differential equation $\dot{X} = -\text{grad } F(X)$. Then, the tangent spaces at X_c to the stable manifold $W_s(X_c)$ and to the unstable manifold $W_u(X_c)$ of X_c are expressed, respectively, as*

$$T_{X_c} W_s(X_c) = \text{span} \{\xi_{kl}\}_{(k,l) \in I_s}, \quad T_{X_c} W_u(X_c) = \text{span} \{\xi_{kl}\}_{(k,l) \in I_u}, \quad (4.9)$$

where $\{\xi_{kl}\}$ denotes the basis of the tangent space given in Prop. 2.1. Moreover, the dimensions of them are

$$\dim(W_s(X_c)) = \dim(T_{X_c} W_s(X_c)) = \#I_s, \quad \dim(W_u(X_c)) = \dim(T_{X_c} W_u(X_c)) = \#I_u. \quad (4.10)$$

Proof Since X_c is a critical point of F , there exists a $Y_\perp \in \text{St}(n-p, n)$ such that

$$AY_\perp = Y_\perp \text{diag}(\lambda_{i_{p+1}}, \dots, \lambda_{i_n}), \quad Y^T Y_\perp = 0. \quad (4.11)$$

We first note that tangent vectors $\xi = \text{sym}(Y D_1 Y_\perp^T)$, $\eta = \text{sym}(Y D_2 Y_\perp^T) \in T_{X_c} \text{Grass}(p, n)$ with $D_1, D_2 \in \mathbb{R}^{p \times (n-p)}$ are paired to give the inner product of the form

$$\langle \xi, \eta \rangle_{X_c} = \text{tr}(\text{sym}(Y D_1 Y_\perp^T) \text{sym}(Y D_2 Y_\perp^T)) = \frac{1}{2} \text{tr}(D_1^T D_2). \quad (4.12)$$

Further, the Hessian form given in Eq. (3.54) holds for any critical point X_c . Therefore, we conclude that the eigenvalues of the Hessian $\text{Hess } F(X_c)$ at X_c are $\nu_{kl} := (\lambda_{i_{p+l}} - \lambda_{i_k})/2$, $k = 1, \dots, p, l = 1, \dots, n-p$, and the eigenvector associated with ν_{kl} is $\xi_{kl} := \text{sym}(Y E_{kl} Y_\perp^T)$. In fact, we can verify that

$$\begin{aligned} \text{Hess } F(X_c)[\xi_{kl}] &= \text{sym}(\xi_{kl}(A - AX_c - X_c A)) \\ &= \text{sym}\left(\frac{1}{2}(Y E_{kl} Y_\perp^T + Y_\perp E_{kl}^T Y^T)A(I - 2X_c)\right) \end{aligned}$$

$$\begin{aligned} &= \text{sym} \left(\frac{1}{2} (\lambda_{i_{p+l}} Y E_{kl} Y_{\perp}^T + \lambda_{i_k} Y_{\perp} E_{kl}^T Y^T) (I - 2Y Y^T) \right) \\ &= \nu_{kl} \xi_{kl}, \end{aligned} \quad (4.13)$$

where use has been made of $AX_c = X_c A$ and the relation (4.11). The eigenvalue ν_{kl} is positive if and only if $(k, l) \in I_s$ and negative if and only if $(k, l) \in I_u$, which results in the conclusion, (4.9) and (4.10). This ends the proof. \square

If $\lambda_p \neq \lambda_{p+1}$, one has $\dim(W_s(X_c)) < \dim(\text{Grass}(p, n))$ for any critical point X_c of F except for the global optimal solution X_* . This means that each stable manifold is thin (or of measure zero), so that most of the sequences $\{X_k\}$ generated by the steepest descent method does not converge to X_c ($\neq X_*$). Since Thm. 3.1 guarantees that $\{X_k\}$ converges to one of the critical points, $\{X_k\}$ usually converges to the global optimal solution X_* , and hence sufficient iterations of the steepest descent method drive the sequence into the convergence region of Newton's method for X_* .

4.5 Some sufficient conditions for convergence to a saddle point

In this subsection, we discuss sufficient conditions for a sequence to converge to a saddle point by means of Algorithm 3.5 with the QR-based retraction (2.43). However, numerical experiments, which are given in the succeeding subsection, show that the convergence to a saddle point hardly occurs.

The condition stated in the following proposition means that the p -dimensional subspace $W_0 \subset \mathbb{R}^n$ represented by the initial point $X_0 \in \text{Grass}(p, n)$ is perpendicular to one of eigenvectors associated with the first p eigenvalues of A in ascending order.

Proposition 4.4 *Let v_1, \dots, v_n are orthonormal eigenvectors associated with the eigenvalues $\lambda_1 \leq \dots \leq \lambda_p < \lambda_{p+1} \leq \dots \leq \lambda_n$ of A , respectively. If the initial point $X_0 \in \text{Grass}(p, n)$ satisfies $X_0 v_i = 0$ for some $i \in \{1, \dots, p\}$, then the sequence $\{X_k\}$ generated by Algorithm 3.5 with the QR-based retraction (2.43) does not converge to the global optimal solution X_* , and hence converges to a saddle point.*

Proof Since the global optimal solution X_* satisfies $X_* v_i = v_i$, it is sufficient to show that $X_k v_i = 0$ for all $k \in \mathbb{N}$ if $X_0 v_i = 0$. We shall show that $X_k v_i = 0$ implies $X_{k+1} v_i = 0$.

Assume that $X_k v_i = 0$. Let P be an $n \times n$ orthogonal matrix whose column vectors are $v_i, v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n$, that is, $P = (v_i, v_1, v_2, \dots, \hat{v}_i, \dots, v_n)$, where \hat{v}_i means that v_i is missing. Then, A is diagonalized into

$$A = P \Lambda P^T, \quad \Lambda = \text{diag}(\lambda_i, \lambda_1, \lambda_2, \dots, \hat{\lambda}_i, \dots, \lambda_n). \quad (4.14)$$

We take a $Y_k \in \text{St}(p, n)$ such that $X_k = Y_k Y_k^T$, and define $Y \in \text{St}(p, n)$ and $X \in \text{Grass}(p, n)$ by $Y = P^T Y_k$ and $X = Y Y^T$, respectively. Note that X_k and X are related by $X_k = Y_k Y_k^T = (PY)(PY)^T = PXP^T$. From the assumption $X_k v_i = 0$, we have

$$0 = X_k v_i = PXP^T v_i = PXe_1, \quad (4.15)$$

where $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^n$. Thus, we obtain $Xe_1 = 0$, and hence X and Y take the form

$$X = \begin{pmatrix} 0 & 0 \\ 0 & X' \end{pmatrix}, \quad X' \in \text{Grass}(p, n-1), \quad (4.16)$$

$$Y = \begin{pmatrix} 0 \\ Y' \end{pmatrix}, \quad Y' \in \text{St}(p, n-1), \quad (4.17)$$

respectively. Let $\Lambda' := \text{diag}(\lambda_1, \dots, \hat{\lambda}_i, \dots, \lambda_n)$. By the use of X and Y of this form, the gradient $\text{grad } F(X_k)$ at X_k is written out as

$$\begin{aligned} \text{grad } F(X_k) &= \text{sym}(AX_k) - X_k \Lambda X_k \\ &= \text{sym}(P \Lambda P^T P X P^T) - P X P^T P \Lambda P^T P X P^T \\ &= P(\text{sym}(\Lambda X) - X \Lambda X) P^T \\ &= P \begin{pmatrix} 0 & 0 \\ 0 & \text{sym}(\Lambda' X') - X' \Lambda' X' \end{pmatrix} P^T. \end{aligned} \quad (4.18)$$

Let η_k denote the search direction $-\text{grad } F(X_k)$ and $t_k > 0$ the step size. Setting $\eta := -(\text{sym}(\Lambda X) - X \Lambda X)$, we obtain $\eta_k = P \eta P^T$. The $\text{qf}((I + t_k \eta_k) Y_k)$ is then written out as

$$\begin{aligned} \text{qf}((I + t_k \eta_k) Y_k) &= \text{qf}(P(I + t_k \eta) Y) \\ &= P \text{qf}((I + t_k \eta) Y) \\ &= P \text{qf} \left(\begin{pmatrix} 1 & 0 \\ 0 & I_{n-1} + t_k (X' \Lambda' X' - \text{sym}(\Lambda' X')) \end{pmatrix} \begin{pmatrix} 0 \\ Y' \end{pmatrix} \right) \\ &= P \begin{pmatrix} 0 \\ * \end{pmatrix}. \end{aligned} \quad (4.19)$$

The succeeding point X_{k+1} is then determined and computed as

$$\begin{aligned} X_{k+1} &= R_{X_k}(t_k \eta_k) \\ &= \text{qf}((I + t_k \eta_k) Y_k) (\text{qf}((I + t_k \eta_k) Y_k))^T \\ &= P \begin{pmatrix} 0 \\ * \end{pmatrix} (0 \ *)^T P^T \\ &= P \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix} P^T. \end{aligned} \quad (4.20)$$

Thus, we have

$$X_{k+1} v_i = P \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix} P^T v_i = P \begin{pmatrix} 0 & 0 \\ 0 & * \end{pmatrix} e_1 = 0. \quad (4.21)$$

This completes the proof. \square

Remark 4.1 This proposition is easily extended to the case where $X_0 v_{i_1} = \dots = X_0 v_{i_l} = 0$ with $i_1, \dots, i_l \in \{1, \dots, p\}$ and $l < p$. In a similar manner to that used in the proof of Prop. 4.4, we can show that the condition $X_0 v_{i_1} = \dots = X_0 v_{i_l} = 0$ implies that $X_k v_{i_1} = \dots = X_k v_{i_l} = 0$ for all $k \in \mathbb{N}$.

There is another condition for convergence to a saddle point. In this condition concerning the choice of an initial point, the original problem is naturally broken up into subproblems of smaller size.

Proposition 4.5 *Let v_1, \dots, v_n are orthonormal eigenvectors associated with the eigenvalues $\lambda_1 \leq \dots \leq \lambda_p < \lambda_{p+1} \leq \dots \leq \lambda_n$ of A , respectively. Let $P := (v_1, \dots, v_n) \in O(n)$. Assume that for an initial point $X_0 \in \text{Grass}(p, n)$, there exist $X_0^{(1)} \in \text{Grass}(p_1, n_1)$ and $X_0^{(2)} \in \text{Grass}(p_2, n_2)$ such that*

$$P^T X_0 P = \begin{pmatrix} X_0^{(1)} & 0 \\ 0 & X_0^{(2)} \end{pmatrix}, \quad (4.22)$$

where $n_1 \geq p$ and $p_1 < p$. Then, the sequence $\{X_k\}$ generated by Algorithm 3.5 with the QR-based retraction (2.43) does not converge to the global optimal solution X_* , and hence converges to a saddle point.

Proof We show that $P^T X_k P = \begin{pmatrix} X_k^{(1)} & 0 \\ 0 & X_k^{(2)} \end{pmatrix}$ for all $k \in \mathbb{N}$, where $X_k^{(1)} \in \text{Grass}(p_1, n_1)$ and $X_k^{(2)} \in \text{Grass}(p_2, n_2)$, if X_0 takes the form (4.22). Then, $\{X_k\}$ should converge to X_∞ such that $P^T X_\infty P = \begin{pmatrix} X_\infty^{(1)} & 0 \\ 0 & X_\infty^{(2)} \end{pmatrix}$ with $X_\infty^{(1)} \in \text{Grass}(p_1, n_1)$ and $X_\infty^{(2)} \in \text{Grass}(p_2, n_2)$. The target point is not the global optimal solution, for which one has $P^T X_* P = \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}$ with $n_1 \geq p$ and $p_1 < p$. We now show that if X_k takes the form

$$P^T X_k P = \begin{pmatrix} X^{(1)} & 0 \\ 0 & X^{(2)} \end{pmatrix}, \quad X^{(1)} \in \text{Grass}(p_1, n_1), \quad X^{(2)} \in \text{Grass}(p_2, n_2), \quad (4.23)$$

then X_{k+1} takes the same form

$$P^T X_{k+1} P = \begin{pmatrix} X_+^{(1)} & 0 \\ 0 & X_+^{(2)} \end{pmatrix}, \quad X_+^{(1)} \in \text{Grass}(p_1, n_1), \quad X_+^{(2)} \in \text{Grass}(p_2, n_2). \quad (4.24)$$

Assume that (4.23) holds. According to the block diagonalization of the right-hand side of (4.23), we put A in the form

$$A = P \Lambda P^T = P \begin{pmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{pmatrix} P^T, \quad (4.25)$$

where $\Lambda := \text{diag}(\lambda_1, \dots, \lambda_n)$, $\Lambda_1 := \text{diag}(\lambda_1, \dots, \lambda_p)$, and $\Lambda_2 := \text{diag}(\lambda_{p+1}, \dots, \lambda_n)$. Setting $X := P^T X_k P$, $\eta := -(\text{sym}(\Lambda X) - X \Lambda X)$, $\eta^{(1)} := -(\text{sym}(\Lambda_1 X^{(1)}) - X^{(1)} \Lambda_1 X^{(1)})$, and $\eta^{(2)} := -(\text{sym}(\Lambda_2 X^{(2)}) - X^{(2)} \Lambda_2 X^{(2)})$, as in the course of the proof of Prop. 4.4, we can write out the search direction $\eta_k := -\text{grad } F(X_k)$ as

$$\eta_k = P \eta P^T = P \begin{pmatrix} \eta^{(1)} & 0 \\ 0 & \eta^{(2)} \end{pmatrix} P^T. \quad (4.26)$$

Let $Y^{(1)} \in \text{St}(p_1, n_1)$ and $Y^{(2)} \in \text{St}(p_2, n_2)$ be matrices satisfying $X^{(1)} = Y^{(1)} (Y^{(1)})^T$ and $X^{(2)} = Y^{(2)} (Y^{(2)})^T$, respectively, and set $Y := \begin{pmatrix} Y^{(1)} & 0 \\ 0 & Y^{(2)} \end{pmatrix} \in \text{St}(p, n)$ and $Y_k := PY$. Then we have $X = YY^T$ and $X_k = Y_k Y_k^T$. Using these Y_k , Y , and the step size $t_k > 0$, we compute the $\text{qf}((I + t_k \eta_k)Y_k)$ to find that

$$\begin{aligned} \text{qf}((I + t_k \eta_k)Y_k) &= P \text{qf}((I + t_k \eta)Y) \\ &= P \text{qf} \begin{pmatrix} (I_{n_1} + t_k \eta^{(1)})Y^{(1)} & 0 \\ 0 & (I_{n_2} + t_k \eta^{(2)})Y^{(2)} \end{pmatrix} \\ &= P \begin{pmatrix} \text{qf}((I_{n_1} + t_k \eta^{(1)})Y^{(1)}) & 0 \\ 0 & \text{qf}((I_{n_2} + t_k \eta^{(2)})Y^{(2)}) \end{pmatrix}. \end{aligned} \quad (4.27)$$

From this expression of $\text{qf}((I + t_k \eta_k)Y_k)$, the succeeding point X_{k+1} is determined and expressed as

$$\begin{aligned} X_{k+1} &= R_{X_k}(t_k \eta_k) \\ &= \text{qf}((I + t_k \eta_k)Y_k) (\text{qf}((I + t_k \eta_k)Y_k))^T \\ &= P \begin{pmatrix} R_{X^{(1)}}(t_k \eta^{(1)}) & 0 \\ 0 & R_{X^{(2)}}(t_k \eta^{(2)}) \end{pmatrix} P^T. \end{aligned} \quad (4.28)$$

Since $R_{X^{(1)}}(t_k \eta^{(1)}) \in \text{Grass}(p_1, n_1)$ and $R_{X^{(2)}}(t_k \eta^{(2)}) \in \text{Grass}(p_2, n_2)$, Eq. (4.28) comes to Eq. (4.24) on setting $X_+^{(1)} = R_{X^{(1)}}(t_k \eta^{(1)})$ and $X_+^{(2)} = R_{X^{(2)}}(t_k \eta_k^{(2)})$. This completes the proof. \square

4.6 Numerical verification of global convergence

Theoretically, there exist sequences which converge to some of saddle points. However, we need not to care about such sequences, practically. In fact, even if the condition $X_0 v_i = 0$, which is referred to in Prop. 4.4, is satisfied for this initial point $X_0 \in \text{Grass}(p, n)$ within the computer accuracy, the computational error allows the sequence to converge to the global optimal solution. We observe this in the following experiment.

We take the case $n = 10, p = 5$. By using a randomly chosen 10×10 orthogonal matrix $P = (v_1, \dots, v_{10})$ and the 10×10 diagonal matrix $A := \text{diag}(1, \dots, 10)$, we define a symmetric matrix by $A := PAP^T$. We next determine $Y_0 \in \text{St}(5, 10)$ by $Y_0 := (v_3, v_4, \dots, v_{10})Q$, where Q is a randomly chosen 8×5 orthonormal matrix, so that each column of Y_0 is a linear combination of v_3, v_4, \dots, v_{10} . The corresponding $X_0 \in \text{Grass}(5, 10)$, which is defined as $X_0 := Y_0 Y_0^T$, satisfies $X_0 v_1 = X_0 v_2 = 0$. It follows from Rem. 4.1 that the sequence $\{X_k\}$ generated by Algorithm 3.5 with this initial point X_0 satisfies $X_k v_1 = X_k v_2 = 0$ for all $k \in \mathbb{N}$. Theoretically, we then expect the sequence to converge to the saddle point $X_{c_1} = Y_{c_1} Y_{c_1}^T$ with $Y_{c_1} = (v_3, v_4, \dots, v_7) \in \text{St}(5, 10)$, at which the objective function F is evaluated as 12.5. Our numerical experiment is shown in Fig. 4.4. We observe that the value of the objective function along $\{X_k\}$ steps down to the minimum. The sequence seems to converge to X_{c_1} at first, which is seen until about thirty iterations.

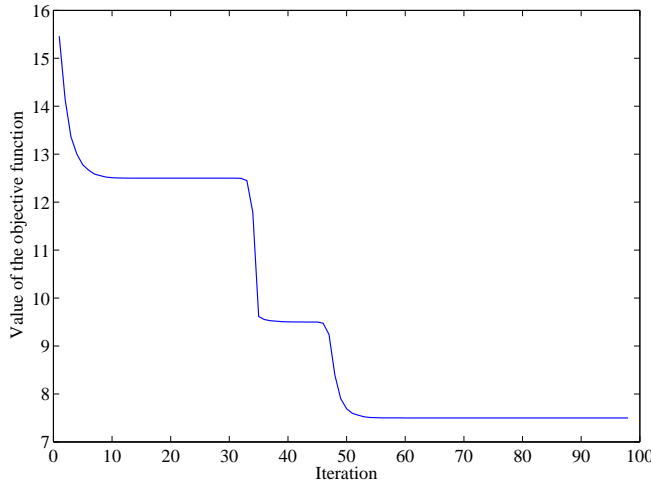


Fig. 4.4 $n = 10$, $p = 5$, and the subspace represented by $X_0 \in \text{Grass}(5, 10)$ is perpendicular to two eigenvectors associated with the two smallest eigenvalues of A .

However, the computing error actually violates the condition $X_k v_1 = 0$ which holds in theory, after about thirty iterations, and the sequence in turn approaches the saddle point $X_{c_2} = Y_{c_2} Y_{c_2}^T$ with $Y_{c_2} = (v_1, v_3, v_4, v_5, v_6)$, at which the value of F is 9.5. Eventually, the condition $X_k v_2 = 0$ is also practically violated to allow the sequence to converge to the global optimal solution $X_* = Y_* Y_*^T$ with $Y_* = (v_1, \dots, v_5)$. The optimal value is $F(X_*) = 7.5$.

We proceed to a numerical experiment concerning Prop. 4.5. Theoretically, Eq. (4.23) implies Eq. (4.24). However, this implication fails actually because of the limitation of computer accuracy. We take the case $n = 10$, $p = 5$, $n_1 = n_2 = 5$, $p_1 = 3$, and $p_2 = 2$. A symmetric matrix A is computed in the same way as in the previous experiment, that is, $A := P \Lambda P^T$, where $P \in O(10)$ and $\Lambda := \text{diag}(1, \dots, 10)$. We next prepare $Y_0^{(1)} \in \text{St}(3, 5)$ and $Y_0^{(2)} \in \text{St}(2, 5)$ with randomly chosen column vectors, and set $Y_0 := P \begin{pmatrix} Y_0^{(1)} & 0 \\ 0 & Y_0^{(2)} \end{pmatrix}$, $X_0^{(1)} := Y_0^{(1)} (Y_0^{(1)})^T$, $X_0^{(2)} := Y_0^{(2)} (Y_0^{(2)})^T$, and $X_0 := P \begin{pmatrix} X_0^{(1)} & 0 \\ 0 & X_0^{(2)} \end{pmatrix} P^T (= Y_0 Y_0^T)$. The X_0 thus determined satisfies the condition (4.22). From Prop. 4.5 and its course of the proof, the sequence $\{X_k\}$ generated by Algorithm 3.5 with X_0 is expected to converge to $X_{c_1} := P \begin{pmatrix} X_{c_1}^{(1)} & 0 \\ 0 & X_{c_1}^{(2)} \end{pmatrix} P^T$, where $X_{c_1}^{(1)} := \begin{pmatrix} I_3 & 0 \\ 0 & 0_2 \end{pmatrix}$ and $X_{c_1}^{(2)} := \begin{pmatrix} I_2 & 0 \\ 0 & 0_3 \end{pmatrix}$. The objective function $F(X)$ then takes the value $F(X_{c_1}) = (1 + 2 + 3 + 6 + 7)/2 = 9.5$ at X_{c_1} . A numerical experiment is shown in Fig. 4.5. The left figure (a) of Fig. 4.5 gives the variation of the values of the objective function $F(X_k)$ along $\{X_k\}$ and the right figure (b) that of the values of $\text{tr}(X_k^{(1)})$, where

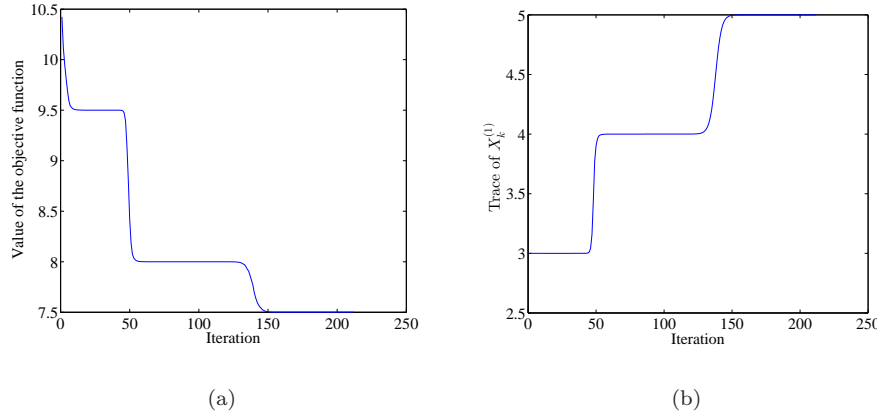


Fig. 4.5 $n = 10$, $p = 5$, and the initial point X_0 reduces the problem into two subproblems at first. Vertical axes take the values of $F(X_k)$ and $\text{tr}(X_k^{(1)})$ along the sequence $\{X_k\}$, respectively.

$X_k^{(1)} \in \mathbb{R}^{n_1 \times n_1}$ is the upper left block of $P^T X_k P$. In the early steps of iterations, the sequence approaches X_{c_1} , as is expected. However, the graph after about fifty iterates shows that the limitation of computer accuracy sets X_k to take the form

$$X_k = P \begin{pmatrix} X_k^{(1)} & 0 \\ 0 & X_k^{(2)} \end{pmatrix} P^T, \quad X_k^{(1)} \in \text{Grass}(4, 5), \quad X_k^{(2)} \in \text{Grass}(1, 5). \quad (4.29)$$

This is confirmed by the fact that the value of $\text{tr}(X_k^{(1)})$ changes from 3 to 4 after about fifty iterations. Once X_k comes down to take the form of Eq. (4.29), the next target of the sequence is $X_{c_2} := P \begin{pmatrix} X_{c_2}^{(1)} & 0 \\ 0 & X_{c_2}^{(2)} \end{pmatrix} P^T$, where $X_{c_2}^{(1)} := \begin{pmatrix} I_4 & 0 \\ 0 & 0 \end{pmatrix}$ and $X_{c_2}^{(2)} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. In fact, $F(X_{c_2}) = (1 + 2 + 3 + 4 + 6)/2 = 8$ is practically reached during the iterations between about 50 and 130, as is seen in Fig. 4.5 (a). After being trapped at intermediate points X_{c_1} and X_{c_2} , the sequence finally converges to the global optimal solution.

These two experiments performed here suggest that the convergence to a saddle point does not occur in practice, and enhance the reliability of our algorithms to find global optimal solutions.

5 Concluding remarks

Treating the geometry of the Grassmann manifold $\text{Grass}(p, n)$ as the set of all orthogonal projection matrices of rank p , we have set up materials necessary for optimization algorithms on $\text{Grass}(p, n)$. In particular, we have worked through the Rayleigh quotient minimization problem on the Grassmann manifold with emphasis on the properties of critical points of the Rayleigh quotient function.

We have addressed the issue encountered in applying Newton's method solely by analyzing the numerical source of the issue. To resolve the issue, we have

composed the steepest descent method and Newton's method to provide a hybrid method, which has been performed for the Rayleigh quotient cost function.

A question as to global convergence is also studied. If an initial point is chosen on the stable manifold of a saddle point with respect to the gradient flow, the generated sequence converges to the saddle point. However, such a manifold is of dimension less than that of the whole space, so that most of sequences converge to a global optimal point. Moreover, even if the initial point is unfortunately put on such a dimension-deficient submanifold, the fluctuation stemmed from the computational error brings the generated sequence out of the submanifold, and allows the sequence to converge to a global optimal solution. Our numerical experiments show that the hybrid method works well in making the convergence speed up. Since Newton's equation for the Rayleigh quotient is expressed as a Lyapunov equation, we can use an existing efficient algorithm for solving Lyapunov equations, which makes our algorithm convenient in practice. We stress that the Lyapunov equation which does not have a unique solution is closely related with the degeneracy of eigenvalues of the symmetric matrix characteristic of the Rayleigh quotient. The set of global optimal solutions is shown to be a submanifold of the Grassmann manifold. This fact is also observed by a numerical experiment. We further point out that our optimization algorithm provides a measure, the objective function, to compare the accuracy among approximate solutions obtained by various existing eigenvalue algorithms. We have shown that our method can provide a best possible solution with respect to this measure.

We make some remarks about previous studies on optimization algorithms. In ordinary optimization methods on the Euclidean spaces, Problem 3.1 is regarded as a problem with constraints. In such a context, a way to solve the problem is to use the augmented Lagrangian method [14,15], which is a kind of a penalty method. However, sequences generated by the augmented Lagrangian method for Problem 3.1 converge very slowly. Moreover, the sequence is likely to converge to a saddle point of the objective function, which is not the global minimum point.

In [12], Helmke, Huper, and Trumpf treated the geometry of the Grassmann manifold in a way different from ours. For example, the geodesic equation (2.28) we obtained in Section 2 is equivalently expressed in [12] as

$$\ddot{X} + [\dot{X}, [\dot{X}, X]] = 0, \quad (5.1)$$

where $[\cdot, \cdot]$ denotes the matrix commutator (Lie bracket). The conclusive resulting algorithms given in [12] is, however, different from our present algorithms. The convergence speed of Newton's method in the present article is faster than that in [12] to the extent of our numerical experiments. This is because Newton's equation in our method is expressed as a Lyapunov equation, to which efficient algorithms can be applied.

A Proof of Proposition 2.4

In this section, we give the proof of Prop. 2.4 on the variational principle.

Proof Since the geodesic equation of a Riemannian manifold is viewed as the equation of motion of a free particle on the Riemannian manifold [5], we consider the Lagrangian L of a

free particle on the Grassmann manifold (2.13), which is put in the form

$$L = \frac{1}{2} \langle \dot{X}, \dot{X} \rangle_X + \text{tr}(\Omega(X^2 - X)) + \lambda(\text{tr}(X) - p), \quad (\text{A.1})$$

where Ω and λ are Lagrange multipliers, and where Ω should be a symmetric matrix on account of the fact that $X^2 - X$ is symmetric. The variation of the Lagrangian L is given by and calculated as

$$\begin{aligned} \delta L &= \text{tr}(\delta \dot{X} \dot{X}) + \text{tr}(\delta \Omega(X^2 - X)) + \text{tr}(\Omega(2X\delta X - \delta X)) + \delta \lambda(\text{tr}(X) - p) + \lambda \text{tr}(\delta X) \\ &= \frac{d}{dt}(\text{tr}(\delta X \dot{X})) + \text{tr}(\delta X(-\ddot{X} + \Omega(2X - I_n) + \lambda I_n)) + \text{tr}(\delta \Omega(X^2 - X)) \\ &\quad + \delta \lambda(\text{tr}(X) - p). \end{aligned} \quad (\text{A.2})$$

On the variational principle, we have

$$\int_{t_1}^{t_2} \delta L dt = 0 \quad (\text{A.3})$$

for any δX subject to the condition $\delta X(t_1) = \delta X(t_2) = 0$. From (A.2) and (A.3), we obtain

$$-\ddot{X} + \Omega(2X - I_n) + \lambda I_n = 0, \quad (\text{A.4})$$

$$X^2 - X = 0, \quad (\text{A.5})$$

$$\text{tr}(X) - p = 0. \quad (\text{A.6})$$

Our next task is to determine Ω and λ . Transposing Eq. (A.4), we have

$$-\ddot{X} + (2X - I_n)\Omega + \lambda I_n = 0. \quad (\text{A.7})$$

Eqs. (A.4) and (A.7) are put together to provide

$$X\Omega = \Omega X. \quad (\text{A.8})$$

Multiplying (A.4) by X from the right and using $X^2 = X$, we obtain

$$\Omega X = \ddot{X}X - \lambda X. \quad (\text{A.9})$$

We see that $\ddot{X}X = X\ddot{X}$ from (A.8) and (A.9). Putting together (A.9) and (A.4), we have

$$\Omega = -\ddot{X} + 2\ddot{X}X + \lambda(I_n - 2X). \quad (\text{A.10})$$

On the other hand, differentiating $X^2 = X$ with respect to t , we obtain

$$\dot{X} = X\dot{X} + 2\dot{X}^2 + \dot{X}X. \quad (\text{A.11})$$

Since $\ddot{X}X = X\ddot{X}$, the above equation becomes

$$\ddot{X} = 2\ddot{X}X + 2\dot{X}^2. \quad (\text{A.12})$$

On account of (A.12), Eq. (A.10) is put in the form

$$\Omega = -2\dot{X}^2 + \lambda(I_n - 2X). \quad (\text{A.13})$$

Substituting (A.13) for Ω in (A.4), we obtain

$$\ddot{X} + 4\dot{X}^2X - 2\dot{X}^2 = 0. \quad (\text{A.14})$$

Since $\dot{X}X + X\dot{X} = \dot{X}$, one has $\dot{X}^2X + \dot{X}X\dot{X} = \dot{X}^2$, and hence Eq. (A.14) is brought into

$$\ddot{X} + 2\dot{X}^2 - 4\dot{X}X\dot{X} = 0. \quad (\text{A.15})$$

This completes the proof. \square

References

1. R. Abraham and J.E. Marsden: Foundations of Mechanics, Benjamin/Cummings Publishing, Reading Mass. (1978)
2. P.-A. Absil, R. Mahony, and R. Sepulchre.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, NJ (2008)
3. P.-A. Absil, R. Mahony, and R. Sepulchre.: Riemannian geometry of Grassmann manifolds with a view on algorithmic computation. *Acta Appl. Math.*, **80**(2):199–220 (2004)
4. R. L. Adler, J. P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub.: Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.*, **22**(3):359–390 (2002)
5. V. I. Arnold, Mathematical Methods of Classical Mechanics. Springer-Verlag, New York, second edition (1989)
6. A. Edelman, T. A. Arias, and S. T. Smith.: The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, **20**(2):303–353 (1998)
7. J. Ferrer, M. I. García, and F. Puerta.: Differentiable families of subspaces. *Linear Algebra Appl.*, **199**:229–252 (1994)
8. K. Fujii.: Note on coherent states and adiabatic connections, curvatures. *J. Math. Phys.*, **41**, 4406–4412 (2000)
9. Z. Gajic and M. T. J. Qureshi.: Lyapunov Matrix Equation in System Stability and Control. Academic Press, Inc (1995)
10. G. H. Golub and C. F. V. Loan.: Matrix Computations. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, third edition (1996)
11. U. Helmke and J. B. Moore.: Optimization and Dynamical Systems. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, With a foreword by R. Brockett (1994)
12. U. Helmke, K. Huper, J. Trumpf.: Newton's method on Grassmann manifolds. *arXiv: 0709.2205v2* (2007)
13. T. Kailath, A. H. Sayed, B. Hassibi.: Linear estimation. Prentice Hall (2000)
14. J. Nocedal and S. J. Wright.: Numerical Optimization. Springer Series in Operations Research. Springer-Verlag, New York (1999)
15. J. A. Snyman.: Practical mathematical optimization: an introduction to basic optimization theory and classical and new gradient-based algorithms. Springer Verlag (2005)
16. S. Tanimura, M. Nakahara, and D. Hayashi.: Exact solutions of the isoholonomic problem and the optimal control problem in holonomic quantum computation. *J. Math. Phys.* **46**, 022101 (2005)
17. L. N. Trefethen and D. Bau.: Numerical Linear Algebra. SIAM (1997)
18. Y.-C. Wong.: Differential geometry of Grassmann manifolds. *Proc. Natl. Acad. Sci. USA*, **57**, 589–594 (1967)